



Analyse fonctionnelle de TaGW2, une E3 ligase de type RING, dans le développement du grain de blé tendre (*Triticum aestivum*)

Julie Bednarek

► To cite this version:

Julie Bednarek. Analyse fonctionnelle de TaGW2, une E3 ligase de type RING, dans le développement du grain de blé tendre (*Triticum aestivum*). Sciences agricoles. Université Blaise Pascal - Clermont-Ferrand II, 2012. Français. NNT : 2012CLF22309 . tel-00857341

HAL Id: tel-00857341

<https://theses.hal.science/tel-00857341>

Submitted on 3 Sep 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N° D. U. : 2309



ANNEE 2012

ECOLE DOCTORALE
SCIENCES DE LA VIE, SANTE, AGRONOMIE, ENVIRONNEMENT
N° d'ordre : 593

Thèse

Présentée à l'Université Blaise Pascal

Pour l'obtention du grade de

DOCTEUR D'UNIVERSITE

Spécialité : Physiologie et génétique moléculaire

Soutenue le 7 Décembre 2012

Julie BEDNAREK

Analyse fonctionnelle de TaGW2, une E3 ligase de type RING, dans le développement du grain de blé tendre (*Triticum aestivum*)

Président : Mr. Didier ATTAIX, Directeur de recherche, INRA - Clermont-Ferrand

Rapporteurs : Mr. Emmanuel GUIDERDONI, Directeur de recherche, CIRAD - Montpellier
Mr. Dominique JOB, Directeur de recherche, CNRS/Bayer- Lyon

Membres : Mr. Jean MOLINIER, Chargé de recherche – CNRS - Strasbourg (Examinateur)
Mr. Said MOUZEYAR, Professeur, Université Blaise Pascal (Directeur de thèse)
Mr. M. Fouad BOUZIDI, Professeur, Université Blaise Pascal (Co-Directeur de thèse)



Université Blaise Pascal – INRA 24, avenue des Landais 63171 Aubière Cedex

Je tiens tout d'abord à remercier les membres de mon jury de thèse, Messieurs Emmanuel GUIDERDONI et Dominique JOB, rapporteurs de cette thèse, et Messieurs Didier ATTAIX et Jean MOLINIER, examinateurs de cette thèse, d'avoir accepté de juger mon travail.

Je remercie également les membres de mon comité de pilotage de thèse, Messieurs Jean MOLINIER, Jacques ROUSTER et Jérôme SALSE, d'avoir évalué l'avancement des travaux engagés au début de cette thèse.

Je tiens à remercier Gilles CHARMET et Thierry LANGIN, de m'avoir accueillie au sein de l'UMR 1095 Génétique, Diversité et Ecophysiologie des Céréales (GDEC). Je remercie tout particulièrement Gilles CHARMET pour l'intérêt constant qu'il a porté à l'avancement de mon travail de thèse au cours de ces années et pour sa disponibilité.

Je remercie Jacques Le Gouis de m'avoir accueillie au sein de l'équipe « Rendement et Adaptation du blé aux contraintes abiotiques (ABC) ». Je te remercie pour ta disponibilité et ton aide précieuse sur les travaux de cartographie.

Je tiens à remercier les Professeurs Said MOUZEYAR et Fouad BOUZIDI pour la confiance et la liberté qu'ils m'ont accordées pour mener à bien ces travaux de thèse. Je vous suis très reconnaissante de m'avoir formée avec beaucoup d'intérêt aux métiers de chercheur et d'enseignant. Je vous remercie pour tout le temps que vous avez pu consacrer aux nombreuses discussions scientifiques (... et autres !) mais également pour votre soutien moral. Je souhaite à tous les thésards un tel cadre de travail.

Du fait de la multidisciplinarité de ces travaux de thèse, j'ai été amenée à travailler avec de nombreuses personnes. Ce travail est plus particulièrement le fruit d'une étroite collaboration avec :

Catherine RAVEL, je te remercie chaleureusement pour toute l'aide que tu m'as apportée aux analyses de séquences, au développement de marqueurs et aux analyses d'association. Je te suis également très reconnaissante de m'avoir formée à la rédaction d'article scientifique. Tu as toujours consacré beaucoup de temps à ce projet et tu as été un réel soutien professionnel et personnel. Un grand MERCI !

Christine GIROUSSE, je te remercie beaucoup pour ton aide précieuse sur les travaux de transgénèse et l'écriture de la version finale de l'article. Tu as toujours été très disponible pour moi, pour me guider et me conseiller, et je te suis très redevable du grand nombre d'heures que tu as dû consacrer au phénotypage des lignées transgéniques. Ta confiance m'a beaucoup aidé.

Je tiens également à remercier François BALFOURIER et Florence EXBRAYAT-VINSON, pour la mise à disposition du matériel biologique de la core collection mondiale de blé tendre et les travaux de génotypage. Florence, merci d'avoir consacré autant de temps à mon initiation à l'utilisation des robots de la plateforme. Je remercie Pierre BARRET et Caroline TASSY pour la production des lignées transgéniques et leurs conseils avisés pour cette étude. Je remercie également Pierre SOURDILLE et Séverine ROUGEOL pour les travaux de cartographie ; Nicole ALLARD et Joëlle MESSAOUD pour leur aide précieuse concernant le phénotypage 'taille du grain' des grains transgéniques (un véritable travail d'orfèvre, MERCI !) ; Annie FAYE, Isabelle

NADAUD et Gérard BRANLARD, pour leur aide et conseils sur les moutures et la mise à disposition de tissus du grain, et Mireille DARDEVET pour l'assignation chromosomique des copies. Je tiens également à remercier Krystof JAGLA et Caroline VACHIAS pour leur aide et leurs conseils pour l'utilisation du microscope confocal sur la plateforme de microscopie de l'UMR 1093 GReD.

Je remercie également Stéphane BENEDIT, Pierre DESRAY et Vincent ALLARD pour les moments de pause agréables passés en leur compagnie lors de mes venues sur le site de Crouël.

Parce qu'une partie importante de ces dernières années a également été consacrée à des activités d'enseignement, je tiens à remercier les différentes équipes d'enseignement avec lesquelles j'ai collaboré. Merci tout particulièrement à Nathalie LEBLANC, Patricia DREVET, Catherine LENNE, Gérard LEDOIGT et Isabelle BALANDIER. Je remercie chaleureusement Frédéric MOREL d'avoir tout mis en œuvre pour me faciliter mes activités d'enseignement cette dernière année passée, afin que je puisse dégager un maximum de temps pour ce travail de thèse. Un grand MERCI à Sylvaine, Christelle, Céline, Christophe et Adèle pour leur aide technique précieuse au cours de mes séances de TP.

Merci également à Nelly et Christine pour la conduite des plantes en serre ; à Dominique pour ces magnifiques photos et sa disponibilité (tu sublimes les résultats !!) ; à Norbert pour son extrême gentillesse (toujours prêt à rendre service !) ; à Sandrine, notre secrétaire en or (merci d'avoir toujours le sourire !) ; et à Jean-Pierre, merci pour ces innombrables discussions sur le temps de midi qui font du bien!

J'ai une pensée toute particulière pour tous les anciens et actuels membres de mon équipe à l'Université. Un énorme MERCI à Jane, Aurélia, Patrick, Monique, Caroline, Imen, Senda, Delphine et Florence. Aurélia, je te remercie pour toute ton aide et tes conseils avisés, notamment pour les travaux de microscopie; Jane, merci du fond du cœur pour ton soutien moral ; Florence, je te remercie de m'avoir formée aux techniques de biologie moléculaire quand je suis arrivée ; Patrick, merci pour ton humour décalé qui m'a fait décompresser un bon nombre de fois (!) ; Monique ... je ne trouve de mots assez forts pour te montrer tout ma gratitude, MERCI d'avoir été cette aide si précieuse au quotidien, tant au niveau technique que moral. Merci d'être si généreuse et attentionnée. Imen, Senda et Delphine (la girlzz team), merci infiniment pour tous les moments que l'on a passé ensemble dans et hors du labo. Delphine, merci d'avoir ensoleillé le labo (c'est le comble pour une belge quand même ...) au cours de ces années passées en ta compagnie et de m'avoir fait rire si souvent ! Je suis heureuse de t'avoir eu pour collègue et maintenant pour amie. Caro, merci pour ton amitié, ta spontanéité, ta gentillesse et ton engouement grandissant pour la recherche ! Je suis extrêmement contente que tu aies choisi de poursuivre sur une thèse au labo et d'avoir ainsi la possibilité de travailler à tes côtés (et sortir bien sûr !!!) encore pendant quelques mois.

Julie, je te remercie du fond du cœur d'être mon amie. Tu m'as formée, conseillé, soutenue, écouté et surtout obligée bon nombre de fois à sortir la tête de mes expériences. Merci infiniment pour tous les moments géniaux entre amies en off!

Je remercie également chaleureusement Ludo, Aurélien, Julien, Laurent et Priscilla pour tous les délires qu'on a pu partager ensemble et ces innombrables discussions scientifiques (... ou pas). Priscilla et Ludo, merci pour votre amitié de longue date, à toutes épreuves. Vous êtes un soutien sans faille ! Merci également à Imen M.,

Raffik, Johann P., Isabelle, Olivier, Marine D., Saman, Pierre M., Anaïs, Delphine G., Géraldine T., Hugues, David L. et Boris F.

Je tiens sincèrement à remercier mes amis à l'extérieur du labo Ysa, Philip, David F., Marion, Charly et Xavier. Vous m'avez aidée à déconnecter un très grand nombre de fois et je vous en remercie du fond du cœur. Philip, MERCI d'avoir été là toutes ces années ... et d'avoir essayé de comprendre (en vain) ce que je faisais !!

Je tiens profondément à remercier ma famille. Je remercie infiniment mes parents de m'avoir élevée dans l'ouverture d'esprit, de m'avoir toujours tirée vers le haut, encouragée et aimée. Merci d'avoir cru en moi. Je vous suis extrêmement reconnaissante de m'avoir aidée à devenir la personne que je suis. Je remercie également mon frère et ma sœur, Alex et Anne-So, pour les moments de bonheur immense que je vis depuis leur naissance. Merci d'être si uniques et de me faire rire aux larmes si souvent !

Je remercie chaleureusement Annie et Robert, de m'avoir toujours soutenue et aimée comme leur propre fille. Merci également à Romain, Sara et Kelis pour ces moments merveilleux en famille !

Pierre, cette thèse t'est entièrement dédiée. Je ne serai jamais arrivée là sans toi, je te dois tout... Tu m'as toujours poussée à être meilleure, à me surpasser ! Je te suis infiniment reconnaissante des sacrifices que tu as dû faire toutes ces années, et de ton soutien permanent au quotidien. Merci d'être mon âme sœur...

RESUME

Le blé tendre, *Triticum aestivum*, est une des céréales les plus cultivées au monde et est d'une importance considérable pour l'alimentation humaine, fournissant environ un cinquième des calories consommées par l'Homme. Le rendement en grain chez les céréales dépend majoritairement du nombre et de la taille des grains. Chez le riz (*Oryza sativa*), le gène *GW2* a été isolé dans un locus à effet quantitatif majeur pour la taille et le poids du grain. Ce gène code pour une enzyme E3 ligase de type RING, qui régule négativement la taille et le poids du grain de riz.

L'homologue de *GW2* chez le blé tendre, le gène *TaGW2*, est exprimé par trois copies *TaGW2-A*, *TaGW2-B* et *TaGW2-D*, portées par chacun des génomes homéologues A, B et D. Les trois copies présentent des profils d'expression distincts au cours du développement du grain. *TaGW2-A* a été cartographié dans une région de QTLs pour le rendement, sur le chromosome 6AS ; et du polymorphisme dans sa séquence promotrice et intronique a été retrouvé associé au poids de 1000-grains dans une core collection mondiale de blé tendre. Afin de rechercher la fonction de *TaGW2*, l'extinction stable des trois copies *TaGW2* a été entreprise par ARN interférence. De manière surprenante, les plantes transgéniques montrent des réductions significatives des dimensions et du poids du grain de blé (- 22,5 et - 30% du volume et de la masse du grain, respectivement), ainsi que du nombre de cellules de l'albumen (- 25%), comparé aux plantes témoins dans nos conditions ; suggérant que *TaGW2* est un régulateur positif de la taille finale du grain chez le blé tendre. La protéine TaGW2-A a été caractérisée aux niveaux moléculaire et biochimique : elle est une E3 ubiquitine ligase fonctionnelle *in vitro*, et s'accumule dans la cellule au niveau du nucléole, du nucléoplasme et du cytoplasme. Sa fonction E3 ligase semble notamment influencer sa localisation subcellulaire. Afin de déterminer la ou les voie(s) de signalisation dans la(es)quelle(s) intervient TaGW2, une banque ADNc de grains de blé a été construite et criblée par double-hybride avec 320 acides aminés de la protéine TaGW2-A. Les premiers interacteurs potentiels identifiés suggèrent d'une part un rôle de TaGW2 dans la régulation de la division cellulaire, et d'autre part une fonction E3 Nedd8 ligase, en plus de son activité E3 ligase.

Mots-clés : blé tendre *Triticum aestivum*, *TaGW2*, développement du grain, ARN interférence, E3 ligase RING

ABSTRACT

Wheat, *Triticum aestivum*, is one of the world's major cereal crops and is of considerable importance to human nutrition, supplying one-fifth of the calories consumed by humans. For important food crops such as wheat, rice and maize, grain yield mainly depends on grain number and size. In rice (*Oryza sativa*), *GW2* was isolated from a major quantitative trait locus for grain size and weight, and encodes an E3 RING ligase that negatively regulates these yield components.

Wheat has *TaGW2* homologs in A, B and D genomes; and copies show distinct expression pattern during whole grain development in wheat. *TaGW2-A* was mapped in a genomic region on 6AS, encompassing previous reported QTLs for yield; and polymorphisms in *TaGW2-A* (promoter and intron 7) were associated with thousand-grain weight, in a worldwide wheat core collection. To investigate *TaGW2* function, RNA interference was used to down-regulate *TaGW2* transcripts levels. Surprisingly, transgenic wheat lines significantly showed decreased grain weight and size-related dimensions, and endosperm cell number compared to controls. The present study thus suggests that *TaGW2* is a positive regulator of the final grain size in wheat, conversely to *GW2* in rice. Biochemical and molecular analyses of the protein TaGW2-A revealed that 1) TaGW2-A is a functional E3 ubiquitine ligase *in vitro*, 2) TaGW2-A accumulates in the nucleolus, the nucleoplasm, and the cytosol, 3) E3 ubiquitine ligase activity seems to impact TaGW2-A subcellular localization. To investigate the TaGW2 signalling pathway(s), cDNA library from whole wheat grains was built and screened with the bait protein TaGW2(1-320). Preliminary results from the interactomic study suggest that TaGW2 may regulate cell division. Moreover, TaGW2 may also function as an E3 Nedd8 ligase, besides its E3 ubiquitin ligase function.

Keywords: wheat *Triticum aestivum*, *TaGW2*, grain development, RNA interference, E3 RING ligase.

SOMMAIRE

Liste des Abréviations	1
SYNTHESE BIBLIOGRAPHIQUE	3
1. Le blé	4
1.1 Origines du blé	5
1.2 Anatomie du blé	6
1.3 Développement du blé	7
2. Le grain du blé	10
2.1 Structure du grain de blé	10
2.2 Composition du grain de blé	13
2.3 Développement du grain de blé	14
2.4 Le statut hydrique du grain au cours de son développement	19
2.5 L'effet des stress abiotiques sur le développement du grain	20
3. Le rendement en grains	22
3.1 Exemples de QTL de rendement en grains identifiés chez le blé	24
3.2 Exemples de gènes de composantes du rendement en grains isolés et caractérisés chez le riz	25
4. L'ubiquitination	27
4.1 L'ubiquitine	27
4.2 La réaction enzymatique d'ubiquitination	28
4.3 Les divers types d'ubiquitination	28
4.4 Les enzymes déubiquitinases (DUBs)	29
4.5 Les protéines Ubiquitin-like (UBLs)	30
5. La voie de protéolyse ciblée Ubiquitine-Protéasome 26S	30
5.1 Le protéasome 26S	31
5.2 Les fonctions de l'UPS	33
6. Les enzymes E3	33
6.1 Les E3 HECT	34
6.2 Les E3 RING	34
6.3 Les E3 U-box	40
6.4 Les CRLs	41
7. Les E3 et le rendement en grain chez les céréales	42
7.1 Les protéines E3 RING OsDIR1, OsDIS1 et TdRF1	42
7.2 Le complexe APC ^{TAD1} ou APC ^{TE}	44
8. Le gène <i>GW2</i>	45
8.1 Les homologues de <i>GW2</i> chez le maïs, <i>Zea mays</i>	47
8.2 Les homologues de <i>GW2</i> chez le blé tendre, <i>Triticum aestivum</i>	48
9. Objectifs du travail de thèse	51
MATERIEL et METHODES	52
Matériel végétal	53
1. Les plantes de blé tendre <i>Triticum aestivum</i>	53
2. Le tabac (<i>Nicotiana tabacum</i>)	55
Les souches bactériennes	56
Les souches de levures <i>Saccharomyces cerevisiae</i>	57
Les vecteurs plasmidiques	59
Méthodes	63

1. Extraction d'acides nucléiques	63
2. Dessin de couples d'amorces	64
3. Amplification PCR	65
4. Clonage dans des vecteurs plasmidiques	68
5. Mutagenèse dirigée	69
6. Transformation de cellules de tabac	69
7. Observations en microscopie confocale	71
8. Test d'autoubiquitination in vitro	72
9. Analyses statistiques	72
10. Extinction in planta de l'expression de <i>TaGW2</i> par ARN interférence	73
11. Sur-expression in planta de <i>TaGW2-A</i>	80
12. La technique de double-hybride	81
12.1 Le double-hybride nucléaire	82
12.2 Le double-hybride cytoplasmique	91
 CHAPITRE I : Caractérisation moléculaire et biochimique de <i>TaGW2</i>	96
1. Obtention des séquences codantes des 3 homéologues de <i>TaGW2</i>	97
2. Caractérisation génétique de <i>TaGW2-A</i> et <i>-B</i>	98
2.1 Développement de marqueurs polymorphes pour <i>TaGW2-A</i> et <i>-B</i>	98
2.2 Cartographie génétique de <i>TaGW2-A</i> et <i>-B</i>	99
2.3 Analyse d'association dans la core collection mondiale de blé tendre	100
3. Analyse transcriptomique de <i>TaGW2</i>	104
4. Caractérisation de la protéine TaGW2-A	106
4.1 Activité E3 ligase in vitro de la protéine TaGW2-A	106
4.2 Localisation subcellulaire de la protéine TaGW2-A	106
4.3 Identification de domaines responsables de la distribution subcellulaire de la protéine TaGW2-A	107
 Conclusions et Discussions du Chapitre I	110
 CHAPITRE II : Recherche de la fonction de <i>TaGW2</i> dans le grain de blé tendre en développement par validation fonctionnelle	119
1. Production des lignées RNAi <i>TaGW2</i>	121
2. Phénotypage des lignées RNAi <i>TaGW2</i>	124
3. Production de lignées de surexpression de <i>TaGW2-A</i>	125
 Conclusions et Discussions du Chapitre II	127
 CHAPITRE III : Recherche des interacteurs protéiques de TaGW2	134
1. Détermination de la séquence de <i>TaGW2</i> à utiliser dans le système double-hybride nucléaire	136
2. Recherche des partenaires protéiques de TaGW2 par Y2H cytoplasmique	137
2.1 Construction de la banque ADNc normalisée	137
2.2 Vérification de la faisabilité de recherche d'IPP de TaGW2 en double-hybride split-ubiquitine	138
2.3 Criblage de la banque ADNc « division » avec la protéine TaGW2-A pleine-longueur, par split-ubiquitine	139

2.4 Validation de la banque ADNc « division » construite dans le système DUALHunter	140
3. Recherche des partenaires protéiques de TaGW2 par Y2H nucléaire	140
3.1 Construction des banques ADNc	140
3.2 Vérification de la non-auto-activation de la protéine chimère DBD-TaGW2(1-320)	140
3.3 Criblage de la banque « division » avec TaGW2-A(1-320)	141
3.4 Analyse des séquences des clones interacteurs potentiels de TaGW2(1-320)	142
 Conclusions et Discussions du Chapitre III	 143
 DISCUSSION GENERALE et PERSPECTIVES	 150
REFERENCES BIBLIOGRAPHIQUES	161
ARTICLE SCIENTIFIQUE	190
ANNEXES	191

Liste des Figures

Figure 1 : Les cinq premières productions céréalières mondiales, en 2010.	4
Figure 2 : Les dix premiers pays producteurs de blé au niveau mondial, en 2010.	4
Figure 3 : Répartition des productions céréalières en France, en 2010.	4
Figure 4 : Origines et expansion du blé vers l'Europe occidentale suivant la diffusion du néolithique.	5
Figure 5 : Relations phylogénétiques entre le blé tendre (<i>T. aestivum</i>) et ses ancêtres sauvages.	5
Figure 6 : Organisation du génome hexaploïde du blé tendre.	6
Figure 7 : Morphologie du blé tendre, <i>Triticum aestivum</i> .	7
Figure 8 : Echelle de Zadoks décrivant le cycle de développement du blé.	8
Figure 9 : Coupe grossie d'un plateau de tallage.	8
Figure 10 : Vue grossie d'un épi en formation, à la fin de l'initiation florale.	8
Figure 11 : Epi de blé tendre en floraison.	9
Figure 12 : Anatomie du grain de blé tendre.	10
Figure 13 : Observation en microscopie électronique à balayage de l'albumen de maïs en développement.	10
Figure 14 : Coupes longitudinale et transversale d'une section de grain de blé dans l'albumen et les enveloppes.	11
Figure 15 : Classification des protéines du grain de la farine du grain de blé : rapprochement entre les classifications d'Osborne et Shewry.	13
Figure 16 : Observation en microscopie électronique à balayage de granules d'amidon de blé tendre.	14
Figure 17 : Evolutions des masses fraîche et sèche de grains de blé tendre, cultivar Hereward.	15
Figure 18 : Représentation schématique du processus de cellularisation de l'albumen.	16
Figure 19 : Développement de l'albumen de maïs de 4 à 20 JAA.	17
Figure 20 : Accumulation des protéines et de l'amidon au cours du développement de l'albumen, dans le blé tendre cultivars Bobwhite et Hereward.	17
Figure 21 : Sections transversales de grain de blé tendre en développement, cultivar Hereward, avant et pendant la phase de remplissage en réserves.	18
Figure 22 : Dynamique de la matière sèche et du contenu en eau des grains basaux d'épillets centraux de blé tendre, cultivar Achill.	19
Figure 23 : Durées de développement de grain de blé de plants cultivés sous trois conditions de température, de l'anthèse jusqu'à maturité.	20
Figure 24 : Notions de caractères simple et complexe.	22
Figure 25 : Schéma des phases de croissance et de développement du blé et périodes d'établissement des différentes composantes du rendement en grain.	23
Figure 26 : Dimensions du grain de blé: longueur, largeur et épaisseur.	23
Figure 27 : Exemples de gènes contrôlant des composantes du rendement isolés et caractérisés fonctionnellement chez le riz.	25
Figure 28 : Structure tridimensionnelle de l'ubiquitine.	27
Figure 29 : Représentation schématique simplifiée de la réaction d'ubiquitination.	28
Figure 30 : Les différentes formes d'ubiquitination et leurs rôles.	29
Figure 31 : Représentation schématique des différentes possibilités de chaînes d'ubiquitine.	29
Figure 32 : Structure tridimensionnelle de protéines Ubiquitin-like.	30
Figure 33 : Organisation et structure du protéasome 26S.	31

Figure 34 : Les différents types de E3 ligases.	34
Figure 35 : Représentation schématique de la structure d'un domaine RING.	35
Figure 36 : Séquences consensus et nombre de chaque type de domaines RING identifiés chez Arabidopsis.	36
Figure 37 : COP1 est un régulateur central de la transduction du signal de réponse de croissance et de développement à la lumière.	37
Figure 38 : Mode d'action de COP1 dans la dégradation de HY5 en réponse à la lumière visible.	37
Figure 39 : Domaines protéiques et protéines interactrices d'AtCOP1.	38
Figure 40 : Modèle d'activation du locus FLC dépendant de FRIGIDA chez Arabidopsis.	39
Figure 41 : Les différentes familles de Cullin-RING ligases (CRLs).	41
Figure 42 : Phénotypes de lignées de riz transgéniques sur-exprimant <i>OsSDIR1</i> comparées à des lignées transformées avec un vecteur contrôle, suite à 4 et 6 jours de stress hydrique.	43
Figure 43 : Modèle de contrôle du tallage chez le riz par le complexe APC ^{TAD1} .	43
Figure 44 : Modèle de contrôle du tallage chez le riz par le complexe APC ^{TE} .	44
Figure 45 : <i>GW2</i> est un régulateur négatif de la largeur et du poids du grain chez le riz.	45
Figure 46 : Les homologues de <i>GW2</i> chez le maïs, <i>ZmGW2-CHR4</i> et <i>ZmGW2-CHR5</i> .	47
Figure 47 : Un polymorphisme dans le promoteur de <i>TaGW2-A</i> est associé avec la taille et au poids du grain dans un panel d'accèsions chinoises de blé tendre.	49
Figure 48 : Un polymorphisme dans l'exon 8 de <i>TaGW2-A</i> est associé à la largeur du grain dans une population F2 en ségrégation issue du croisement Lankaodali (LK) x Chinese Spring (CS).	50
Figure 49 : Schéma représentatif du principe de génotypage des SNP avec des amorces « extension M13 »: exemple du couple d'amorce A-SNPintron7.	65
Figure 50 : Programmes types d'amplification PCR pour le génotypage des SNP à l'aide d'un jeu de couples d'amorces « extension M13 ». Exemple des couples d'amorce A-SNPintron7 (T/C).	66
Figure 51 : Principe du clonage d'un gène d'intérêt par recombinaison homologue (Technologie Gateway, Invitrogen).	68
Figure 52 : Compte-rendu de la production et de la purification de la protéine 6xHis-TaGW2-A par la société Proteogenix.	72
Figure 53 : Principe de l'ARN interférence.	74
Figure 54 : Représentation schématique de la cassette RNAi <i>TaGW2</i> dans le vecteur pSTARGATE modifié.	77
Figure 55 : Phénotypes retrouvés sur des plantes contrôle et transgéniques RNAi en génération F1.	79
Figure 56 : Principe du double-hybride nucléaire. Exemple de la modularité du facteur de transcription GAL4.	82
Figure 57 : Cycle de vie de <i>S. cerevisiae</i> .	83
Figure 58 : Schéma simplifié du protocole expérimental de création d'une banque ADNc en levures Y187, dans le système Matchmaker.	84
Figure 59 : Zygote de <i>S. cerevisiae</i> en formation.	88
Figure 60 : Principe du système double-hybride DUALHunter, basé sur la modularité de l'ubiquitine (split-ubiquitine).	91
Figure 61: Alignement des séquences nucléotides ADNc des homéologues <i>TaGW2-A</i> et <i>-B</i> de <i>T. aestivum</i> cv Récital, et de l'EST CK206178 de <i>T. aestivum</i> cv Chinese	97

Spring.

Figure 62 : Alignement des séquences complètes nucléotidiques ADNc et des séquences complètes primaires protéiques des homéologues <i>TaGW2-A</i> , <i>-B</i> et <i>-D</i> , de <i>T. aestivum</i> cv Récital.	97
Figure 63 : Cartographie génétique de <i>TaGW2-A</i> , sur la population Ct x Cs.	99
Figure 64 : Cartographie génétique de <i>TaGW2-B</i> , sur la population Re x R.	100
Figure 65 : Représentation schématique du déséquilibre de liaison observé entre les deux marqueurs pour la copie A de <i>TaGW2</i> , <i>TaGW2_A_R</i> (-593) et <i>TaGW2_A_Y1061</i> , et la marqueur <i>cfe273</i> , dans la core collection mondiale de blé tendre.	101
Figure 66 : Distribution géographique mondiale des haplotypes pour <i>TaGW2-A</i> et des allèles identifiés au marqueur pour <i>TaGW2-B</i> .	103
Figure 67 : Accumulation du transcrit <i>TaGW2</i> dans différents organes et dans des feuilles prélevées à différents stades des développements végétatif et reproductif de <i>T. aestivum</i> cv Récital.	104
Figure 68 : Taux d'expression relative des transcrits <i>TaGW2-A</i> , <i>-B</i> et <i>-D</i> au cours du développement du grain de blé tendre, cv Récital.	104
Figure 69 : Taux d'expression relative des transcrits <i>TaGW2-A</i> (A), <i>-B</i> (B) et <i>-D</i> (C) au cours de la phase de remplissage du grain, dans des grains entiers de plantes ayant été placées à 20°C et 28°C.	105
Figure 70 : Activité E3 ubiquitine ligase de la protéine <i>TaGW2-A</i> .	106
Figure 71 : Localisation subcellulaire de la protéine <i>TaGW2-A</i> et identification d'un signal NES fonctionnel.	107
Figure 72 : Identification d'un signal NoLS fonctionnel dans <i>TaGW2-A</i> .	109
Figure 73 : La perte de fonction E3 ligase de la protéine <i>TaGW2-A</i> modifie sa localisation subcellulaire.	109
Figure 74: Positions et allèles des deux marqueurs développés dans le promoteur de <i>TaGW2-A</i> , dans l'étude de Su et ses collaborateurs (2011).	113
Figure 75: Le transport nucléo-cytoplasmique actif de protéines à travers les pores de la membrane nucléaire.	114
Figure 76: L'export nucléaire de protéines dépendant du récepteur CRM1.	114
Figure 77 : Alignements des séquences primaires des protéines homéologues <i>TaGW2-A</i> , <i>-B</i> et <i>-D</i> ; et de <i>TaGW2-A</i> chez le blé tendre et <i>GW2</i> chez le riz.	117
Figure 78: Extrait des séquences génomiques de <i>T. aestivum</i> ciblées par la cassette RNAi <i>TaGW2</i> .	121
Figure 79 : Schéma du protocole expérimental de génération des lignées transgéniques RNAi T105 et T111.	122
Figure 80 : Expression endogène relative de <i>TaGW2</i> , toutes copies confondues, dans des plantes T1 issues des plantes-mère 105.2, 105.3 et 111.1, par rapport à des plantes contrôle nulle-ségrégantes (NS).	123
Figure 81 : Expression endogène relative des trois homéologues <i>TaGW2</i> dans des plantes T2 transgéniques et nulle-ségrégantes.	123
Figure 82 : Phénotypes des grains des plantes T2 des lignées transgéniques RNAi T105 et T111, comparés à des grains de plantes contrôle T2 nulle-ségrégantes (NS).	124
Figure 83 : Caractérisation phénotypique des plantes T2 des lignées RNAi T105 et T111.	124
Figure 84 : Correspondance entre les phénotypes des grains des plantes transgéniques RNAi <i>TaGW2</i> et les processus cellulaires déterminant le potentiel de poids du grain lors de la phase de division.	130
Figure 85 : Relation entre le profil d'accumulation en transcrit <i>TaGW2</i> , les	131

phénotypes des grains transgéniques RNAi <i>TaGW2</i> et les processus cellulaires de détermination du potentiel de taille et de poids du grain.	
Figure 86 : Modèle proposé pour le rôle de <i>TaGW2</i> dans la régulation de la taille et du poids du grain.	133
Figure 87 : Schéma simplifié du protocole expérimental de la recherche d'IPP par chromatographie d'affinité.	135
Figure 88 : Détermination de la séquence maximale non-autoactivatrice de TaGW2.	136
Figure 89 : Stades de développement de grains utilisés pour la création des deux banques ADNc pour un criblage double-hybride nucléaire par le système Matchmaker.	140
Figure 90 : Schéma récapitulatif du crible double-hybride de la banque « division » avec TaGW2(1-320) comme protéine appât.	141
Figure 91 : Représentation schématique des principales étapes de la réaction de neddylation/deneddylation.	147
Figure 92: Les divers effets de la neddylation.	148

Liste des Tableaux

Tableau 1 : Distribution histologique des principaux constituants du grain de blé.	12
Tableau 2 : Caractéristiques des différents stades de développement du grain de blé.	15
Tableau 3 : Exemples de QTLs détectés pour différentes composantes du rendement, chez le blé tendre.	24
Tableau 4 : Exemples de la fonction de gènes de la voie UPS, dans le développement d' <i>Arabidopsis thaliana</i> .	33
Tableau 5 : Génotypes des souches de <i>S. cerevisiae</i> utilisées au cours de ce travail de thèse.	57
Tableau 6 : Composition du milieu réactionnel de réaction PCR sur ADN génomique pour le génotypage des SNP. Exemple d'un des deux couples d'amorce allèle spécifique du couple A-SNPintron7.	65
Tableau 7 : Origines et allèles au marqueur TaGW2_A_Y1061 des treize variétés de blé tendre utilisées pour la recherche de diversité nucléotidique dans les copies de <i>TaGW2</i> .	98
Tableau 8 : Analyse d'association entre le polymorphisme aux marqueurs pour <i>TaGW2-A</i> et le marqueur cfe273, et la variation phénotypique de neuf caractères de rendement et de qualité du grain, dans les 372 individus de la core collection mondiale de blé tendre (CC372).	102
Tableau 9 : Analyse d'association entre le polymorphisme au marqueur pour <i>TaGW2-B</i> et la variation phénotypique de huit caractères de rendement et de qualité du grain, dans les 372 individus de la core collection mondiale de blé tendre (CC372).	102
Tableau 10: QTLs pour le rendement en grain et ses composantes reportés sur le chromosome 6AS.	111
Tableau 11 : Analyse d'association multilocale entre le polymorphisme pour <i>TaGW2-A</i> et le marqueur cfe273, et la variation phénotypique pour le poids de 1000-grains et les quantités d'azote et de protéines dans le grain, dans 193 individus de la core collection mondiale de blé tendre.	112
Tableau 12 : Expression endogène relative de <i>TaGW2</i> , toutes copies confondues, dans des plantes T2 transgéniques et nulle-ségrégantes des événements T105 et T111.	123
Tableau 13 : Tests fonctionnels n°1 et 2 de la construction Ost4-TaGW2A-Cub-LexA(VP16) dans le système Split-Ubiquitine DualHunter.	138
Tableau 14 : Test de détermination de la composition du milieu sélectif à utiliser (test n°3) pour le criblage de la banque « division » avec la protéine appât TaGW2, dans le système double-hybride split-ubiquitine DualHunter.	139
Tableau 15 : Détermination de la non-autoactivation des gènes rapporteurs MEL1 et AUR1C, de la protéine appât TaGW2(1-320).	140
Tableau 16 : Résultat de l'analyse blastx après séquençage des 174 clones interacteurs potentiels de TaGW2(1-320).	142
Tableau 17 : Valeurs des largeurs de grain et des taux d'expression relative de <i>TaGW2</i> de 14 variétés chinoises de blé, fournies en Figure 5 de l'étude de Su <i>et al</i> (2011).	154

LISTE DES ABBREVIATIONS

°CJours : Degré Celsius Jours
 3AT : 3-aminotriazol
 AD: Activation domain
 BiFC : Bimolecular Fluorescence Complementation
 CAPS: Cleaved Amplified Polymorphic Sequence
 CC372: core collection mondiale de blé tender (372 variétés)
 cM: centimorgan
 Cub: moitié C-terminale de la split-ubiquitine
 Cv : cultivar
 DAA: Days after anthesis
 DBD: DNA Binding domain
 IAA : Indole-3-acetic acid
 IPP: interactions protéine-protéine
 JAA : Jours après anthèse
 LOD: logarithm of odds
 NES: Nuclear Export Signal
 NLS: Nuclear Localization Signal
 NoLS: Nucleolar Localization Signal
 NS: Nul(le)s-ségréant(s)
 Nub: moitié N-terminale de la split-ubiquitine
 PMG: Poids de 1000-grains
 QTL : Quantitative Trait Locus
 RING: Really INteresting Gene
 RNAi : RNA interference
 SNP: Single Nucleotide Polymorphism
 Ub : Ubiquitine
 UBL : Ubiquitin-like
 UPS: Ubiquitin Proteasome System
 Y2H: Yeast two-hybrid

SYNTHESE BIBLIOGRAPHIQUE

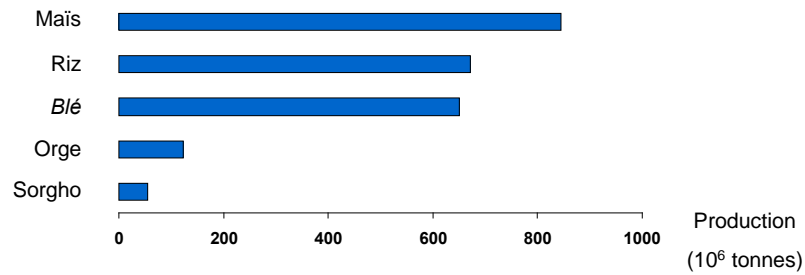


Figure 1 : Les cinq premières productions céréalières mondiales, en 2010. Le blé est la troisième céréale la plus produite dans le monde, derrière le maïs et le riz. Le blé tendre représente 95% de la production mondiale de blé.
Source : FAOstat, <http://faostat.fao.org>

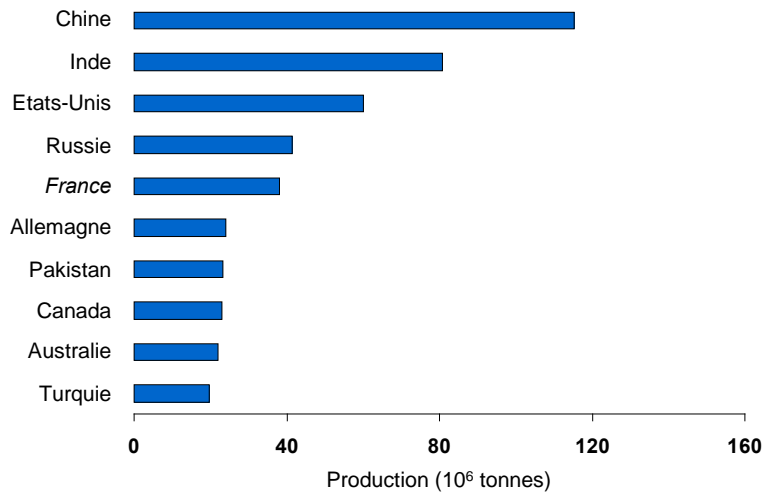


Figure 2 : Les dix premiers pays producteurs de blé au niveau mondial, en 2010. La France est le 5^{ème} producteur mondial de blé, avec une production de 38,2 millions de tonnes.
Source : FAOstat, <http://faostat.fao.org>

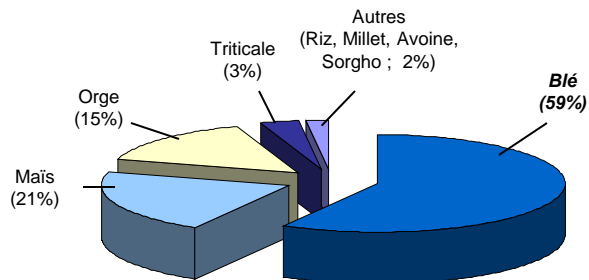


Figure 3 : Répartition des productions céréalières en France, en 2010. Le blé représente la 1^{ère} production céréalière en France.
Source : FAOstat, <http://faostat.fao.org>

1. Le blé

Le blé est une céréale d'importance économique mondiale majeure. Elle est, avec le riz, la céréale la plus consommée par l'Homme et représente un cinquième des apports caloriques dans l'alimentation humaine (Zohary et Hopf, 2000). Avec une production mondiale de 650 millions de tonnes en 2010, le blé est la 3^{ème} céréale la plus produite dans le monde derrière le maïs (844 millions de tonnes) et le riz (672 millions de tonnes ; Figure 1), représentant 26,75% de la production mondiale de céréales ; et constitue la plus grande surface de céréale cultivée dans le monde avec 216 millions d'hectares (source FAOSTAT 2012). Le blé est sans égal quant à l'étendue géographique de sa culture : il est cultivé du 67° parallèle Nord en Scandinavie jusqu'au 45° parallèle Sud en Argentine (Feldman, 1995). Il est également incomparable quant à sa diversité de variétés et la proportion avec laquelle son utilisation est pleinement ancrée dans les cultures et les religions de bon nombre de sociétés diverses (Shewry, 2009).

En France, le blé représente d'autant plus un intérêt économique considérable qu'elle est le 5^{ème} producteur mondial (Figure 2) et le 1^{er} producteur européen dans l'Europe des 27 avec une production de 38,2 millions de tonnes en 2010, soit 19% de la production totale de blé en Europe (source FAOSTAT 2012). En France, le blé représente la 1^{ère} production céréalière devant le maïs (13,9 millions de tonnes) et l'orge (10,1 millions de tonnes ; Figure 3).

Le terme générique 'blé' désigne à la fois les différentes espèces du genre *Triticum* et le grain. Le blé (*Triticum ssp*) est une plante monocotylédone de la famille des Graminées (ou Poacées). C'est une céréale à paille annuelle haute de 0,5 à 1,5m ; cultivée pour sa paille et majoritairement pour ses grains, à inflorescence en épi terminal dense. Les deux espèces les plus cultivées sont le blé tendre (*Triticum aestivum*) et le blé dur (*Triticum durum*), avec une proportion de 95% et 5% de la production de blé mondiale respectivement (Shewry, 2009). Le blé est majoritairement utilisé dans l'alimentation humaine (58%), dans l'alimentation animale (blés fourragers, 34%), mais également à des fins industrielles (8%) dans l'amidonnerie et la glutennerie pour l'industrie cosmétique, la papeterie, et plus récemment la fabrication de matériaux plastiques biodégradables et la production de bioéthanol (Poitrat, 1999).

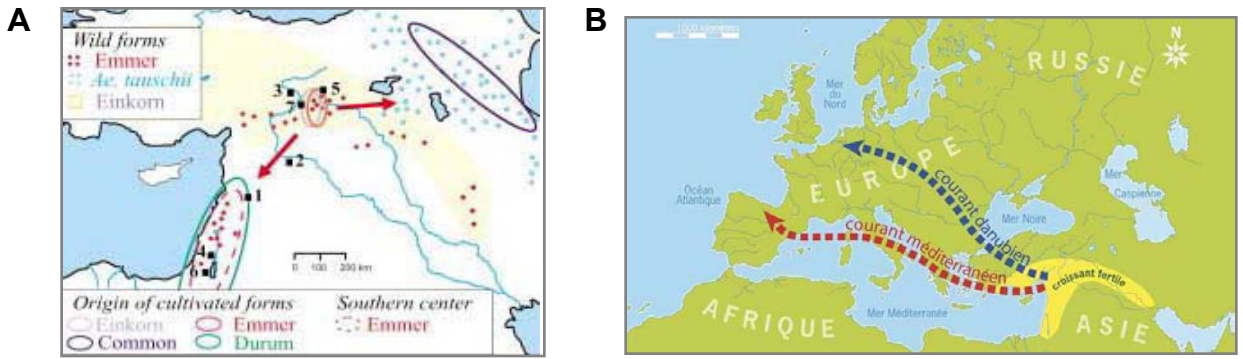


Figure 4 : Origines (A) et expansion du blé vers l'Europe occidentale suivant la diffusion du néolithique (B). (A) Les zones cerclées correspondent aux régions géographiques putatives d'origine des formes cultivées des blés ancestraux. Einkorn, l'engrain ; Emmer, l'amidonner sauvage ; Durum, le blé dur. Les distributions approximatives de l'amidonner sauvage (emmer), d'*Aegilops tauschii* (*Ae. tauschii*) et de l'engrain sauvage (einkorn) sont indiquées par des points rouges et bleus, et la zone jaune, respectivement. Les nombres indiquent les sites archéologiques où des fossiles de céréales domestiquées de plus de 9000 ans ont été retrouvés : 1, Tell Aswad ; 2, Abu Hureyra ; 3, Cafer Höyük ; 4, Jericho ; 5, Cayönü ; 6, Nahal Hemar ; et 7, Nevali Cori. D'après Dubcovsky et Dvorak, 2007; Shewry, 2009. (B) La culture du blé s'est répandue au Néolithique vers l'Europe occidentale au fil des migrations et des échanges commerciaux, suivant deux voies : la voie danubienne au nord et la voie méditerranéenne au sud. (Source INRAP ; et d'après Shewry, 2009).

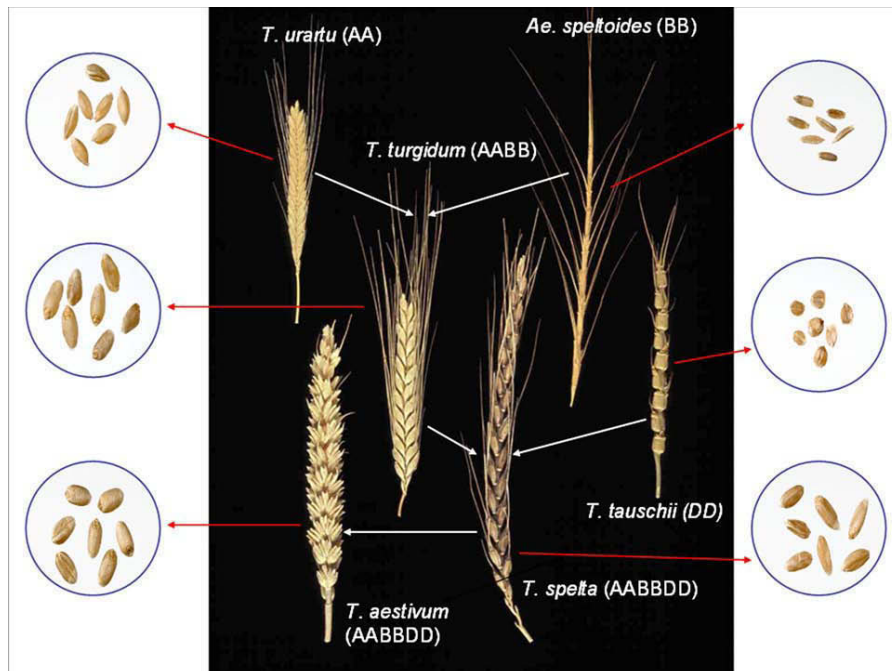


Figure 5 : Relations phylogénétiques entre le blé tendre (*T. aestivum*) et ses ancêtres sauvages (d'après Shewry, 2009). Un exemple d'épi et de grains est donné pour chaque espèce. Le blé tendre résulte de deux hybridations interspécifiques successives. Le croisement entre *Triticum urartu* et une espèce proche de *Aegilops speltoides* a donné *Triticum turgidum*, à l'origine de *Triticum turgidum* ssp *durum*, le blé dur. Le croisement entre *Triticum turgidum* et *Triticum tauschii* a conduit à la formation du blé tendre, *Triticum aestivum*.

Le blé tendre ou froment (*Triticum aestivum*) est cultivé pour sa farine panifiable et est utilisé pour la fabrication d'une large gamme de pains et de produits alimentaires transformés (biscuiterie). La propriété technologique de la pâte fabriquée à partir de farine de blé tendre est unique et a largement contribué au succès du blé tendre : la viscoélasticité de la pâte autorise la fabrication de pains levés en permettant la capture du dioxyde de carbone dégagé lors de la dégradation anaérobie de l'amidon par les levures (levée de la pâte). Le blé dur (*Triticum turgidum* ssp *durum*) est plus adapté au climat sec méditerranéen et est utilisé pour la fabrication de pâtes alimentaires et de semoules.

1.1 Origines du blé tendre

Les blés cultivés sont apparus il y a une dizaine de milliers d'années en Mésopotamie, au moment de la 'révolution néolithique' au cours de laquelle l'Homme est passé d'un mode de vie basé sur la chasse et la cueillette à un mode de vie basé sur l'agriculture et l'élevage, pour subvenir à ses besoins (Shewry, 2009). Le blé est originaire du sud-est de la Turquie, dans la région de Diyarbakir (Figure 4, A). Le blé entra en Europe tout d'abord par l'Anatolie puis la Grèce il y a 8000 ans, puis ensuite par deux voies distinctes il y a 7000 ans (Figure 4, B) : une voie au nord passant par les Balkans et le Danube, et une voie au sud traversant l'Italie, la France et l'Espagne. La culture du blé atteignit le Royaume-Uni et la Scandinavie il y a 5000 ans. Elle s'étendit également à l'Est du Croissant Fertile, par l'Iran, pour atteindre la Chine il y a 3000 ans ; et en Afrique par l'Égypte. Le blé fut introduit au Mexique par les colons espagnols au XVI^e siècle, et en Australie au XVIII^e siècle (Shewry, 2009).

Le genre *Triticum* regroupe des espèces avec des ploïdies variées telles que des espèces diploïdes comme l'engrain (*Triticum monococcum*), des espèces tétraploïdes comme l'amidonniér (*Triticum dicoccum*) ou le blé dur (*Triticum durum*) et des espèces hexaploïdes telles que le blé tendre (*Triticum aestivum*). Le génome actuel du blé tendre provient de 2 événements majeurs d'allopolyploïdisation (Figure 5). Un premier événement de polyploïdisation résultant du croisement entre deux espèces diploïdes ($2n=14$), *Triticum urartu* (AA) et une espèce proche d'*Aegilops sepltoïdes* (BB), a permis l'apparition du blé dur sauvage (*Triticum turgidum* ssp *dicoccoides*) au génome AABB, à l'origine du blé dur actuel *Triticum durum*. Un second croisement entre le tétraploïde *Triticum turgidum* et le diploïde *Triticum tauschii* (aussi appelé *Aegilops tauschii* ou *Aegilops squarosa*, génome DD) a

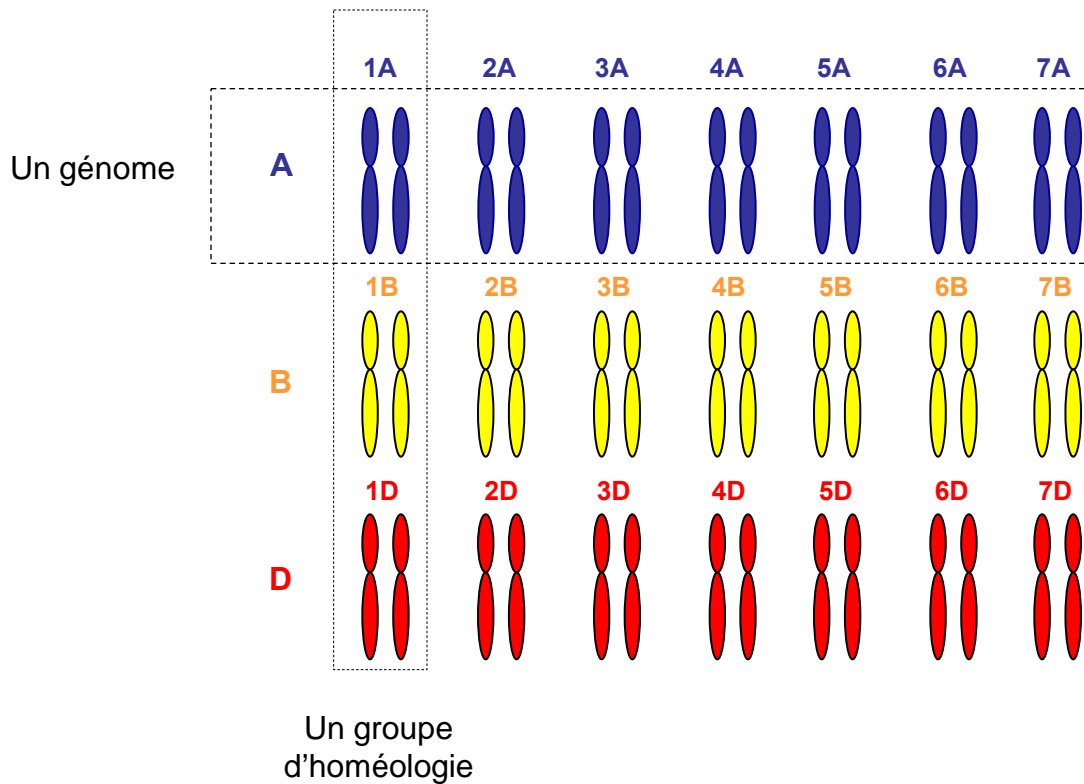


Figure 6 : Organisation du génome hexaploïde du blé tendre. Ce génome comprend 21 chromosomes intégrant les trois génomes homéologues A, B et D, issus de chaque ancêtre. L'ensemble formé par trois chromosomes homéologues est appelé groupe d'homéologie.

conduit à l'obtention de blés hexaploïdes tels que le blé tendre (*Triticum aestivum*) et l'épautre (*Triticum spelta*).

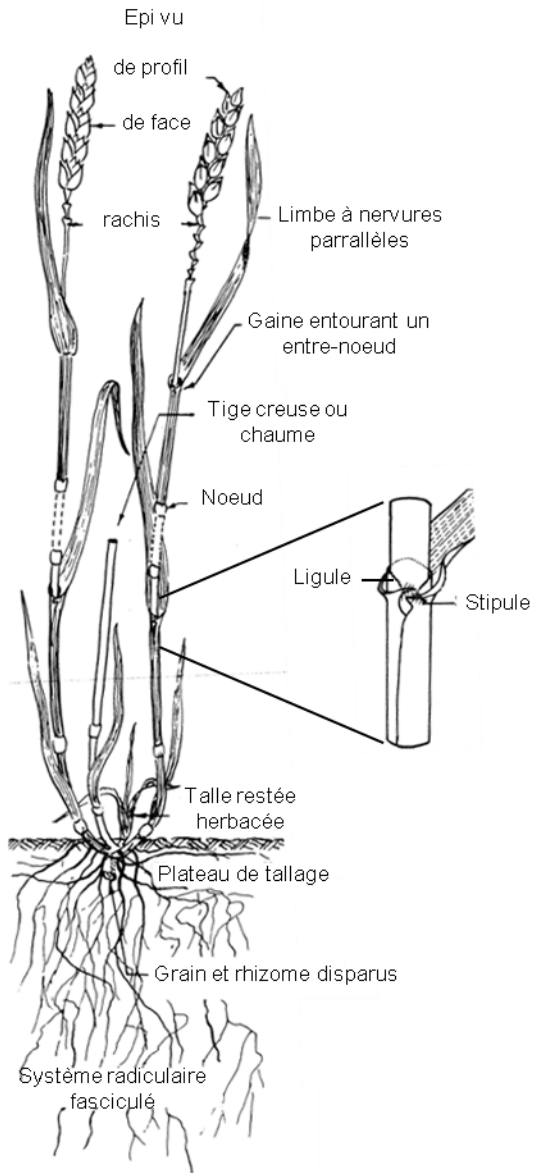
Le génome du blé tendre est structuré en 21 paires de chromosomes intégrant les trois génomes homéologues A, B et D issus de chaque ancêtre (Figure 6). L'ensemble formé par trois paires de chromosomes homéologues se nomme « groupe d'homéologie ». Le génome du blé tendre est de très grande taille, constitué de 17 milliards de paire de bases ce qui représente environ 5 fois la taille du génome humain. Cette taille importante ainsi que la forte proportion de séquences répétées (80% du génome) constituent pour l'instant des obstacles importants à son séquençage complet (Paux *et al*, 2008).

1.2 Anatomie du blé

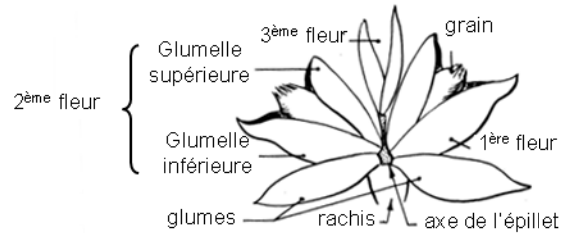
Les espèces du genre *Triticum* sont des herbacées annuelles à feuilles alternes et à croissance définie.

L'appareil végétatif du blé comprend en partie souterraine un système racinaire et en partie aérienne des tiges creuses ou chaumes, portant des feuilles engainantes (Figure 7, A). Le système aérien du blé est formé d'un certain nombre d'unités biologiques appelés talles partantes toutes d'une zone située à la base de la plante, le plateau de tallage. La première talle formée est également appelée 'maître-brin'. Chaque talle, après développement complet de la plante, est formée d'une tige feuillée ou chaume portant à son extrémité une inflorescence. La tige est constituée d'entre-nœuds, séparés par des nœuds, zones méristématiques à partir desquelles, d'une part les entre-nœuds s'allongent et d'autre part les feuilles se différencient. Chaque nœud est le point d'attache d'une feuille. Chaque feuille comprend deux parties : une portion inférieure appelée gaine, enveloppant l'entre-nœud correspondant, et une portion supérieure appelée limbe. Les gaines, attachées au niveau des nœuds, sont emboîtées les unes dans les autres et forment un tube cylindrique entourant la tige qui se déboîte au fur et à mesure de la croissance des entre-nœuds. A la jonction du limbe et de la gaine, on peut trouver une petite membrane non vasculaire, plus ou moins longue et dentelée, la ligule. Chez le blé, de chaque côté de celle-ci, à la base du limbe se trouvent deux stipules plus ou moins embrassantes et velues : les oreillettes (Soltner, 1988).

A – La plante entière



B – L'épillet



C – La fleur

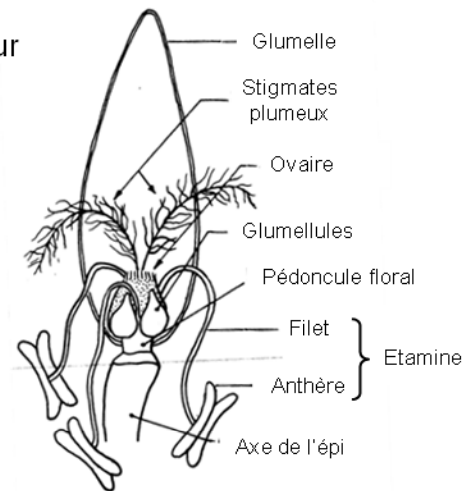


Figure 7 : Morphologie du blé tendre, *Triticum aestivum* (d'après Soltner, 1988). (A) L'appareil végétatif est constitué en partie souterraine d'un système racinaire fasciculé, et en partie aérienne de talles. Chaque talle après complet développement est formée d'une tige feuillée (chaume) portant à son extrémité une inflorescence, l'épi. Les feuilles sont engainantes et issues chacune d'un nœud. L'inflorescence, portée par le rachis, est constituée de 15 à 25 épillets constitués chacun de 3 à 4 fleurs. (B) Un épillet regroupe généralement trois fleurs à l'intérieur de deux glumes. (C) Chaque fleur est entourée de deux glumelles et contient deux pièces stériles, les glumellules, trois étamines et un ovaire surmonté de deux styles plumeux.

Au cours du développement du blé, deux systèmes racinaires vont se succéder : le système de racines primaires ou séminales et le système de racines secondaires ou de tallage. Le système de racines primaires est mis en place de la levée de la graine jusqu'au début du tallage ; le système de racines secondaires est fasciculé et apparaît lorsque la plante commence à émettre les talles secondaires.

L'inflorescence chez le blé est de type «épi», avec une unité morphologique de base qui est l'épillet (Figure 7, B). Le rachis, ou axe de l'épi, porte de 15 à 25 épillets, organisés en 2 rangées situées de part et d'autre de l'axe. L'épillet se compose d'un ensemble de 3 ou 4 fleurs à l'intérieur de deux glumes. Chaque fleur est dépourvue de pétales et est entourée de deux glumelles (inférieure et supérieure ; Figure 7, C). Elle contient trois étamines et un ovaire surmonté de deux styles plumeux. La fleur du blé est dite cléistogame, c'est-à-dire que, le plus souvent, la fécondation a lieu à l'intérieur des glumelles, avant que les étamines n'apparaissent à l'extérieur. De ce fait, l'autofécondation est le mode de reproduction le plus fréquent chez le blé. Les glumes et les glumelles sont éliminées au moment du battage pour libérer le grain.

1.3 Développement du blé

Il existe deux types de blés : les blés d'hiver et les blés de printemps. Les blés d'hiver nécessitent une période de froid prolongée (0 à 7°C, pendant 4 à 8 semaines) appelée vernalisation, pour acquérir l'aptitude à fleurir. Les blés de printemps, quant à eux, ne nécessitent pas de période de vernalisation (Acevedo *et al*, 2002). Les blés tendres d'hiver sont semés à l'automne et caractérisent les régions méditerranéennes et tempérées ; les blés de printemps sont semés au printemps plutôt dans des pays à hiver très rude.

Le cycle de développement du blé se compose tout d'abord d'une période végétative, durant laquelle la plante ne différencie que des feuilles et des racines, puis d'une phase reproductive, dominée par l'apparition de l'épi et la formation du grain. Ce cycle peut être décomposé en stades clés et a été décrit dans les échelles de classification visuelle du développement du blé de Zadoks (Figure 8), de Keller-Baggiolini, et de Feekes-Large (Keller et Baggiolini, 1954 ; Large, 1954 ; Zadoks *et al*, 1974). La durée de chacun des stades de ces phases du développement dépend essentiellement du génotype, de la température, de la

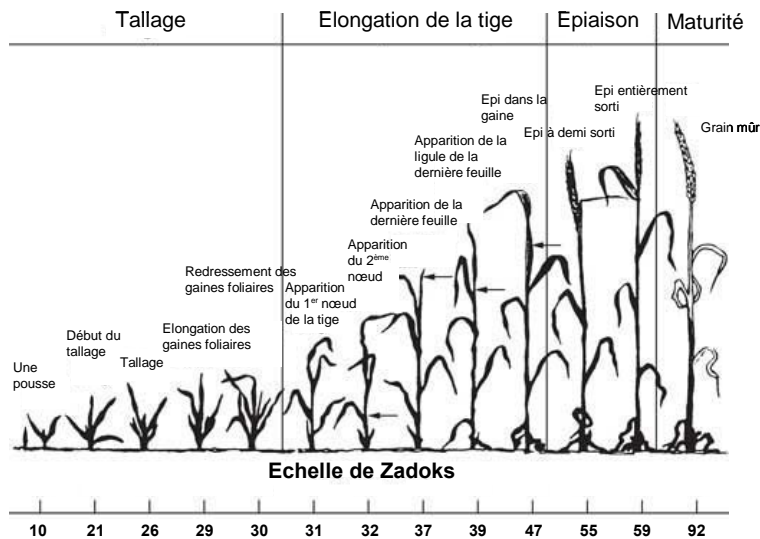


Figure 8 : Echelle de Zadoks décrivant le cycle de développement du blé (d'après Zadoks *et al*, 1974). Chaque étape clé du développement est désignée par un nombre allant de 10 à 92.

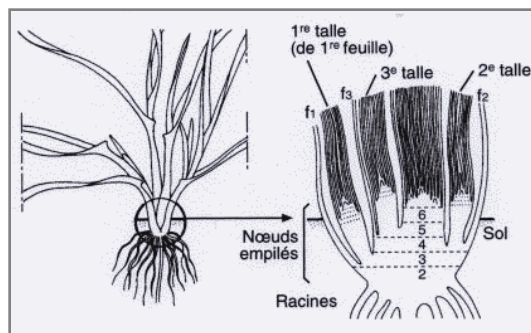


Figure 9 : Coupe grossie d'un plateau de tallage (d'après Boyeldieu, 1997). Sitôt émise la 3^{ème} feuille du jeune plant de blé, le 2^{ème} entre-nœud qui porte le bourgeon terminal s'allonge puis enfle pour former le plateau de tallage. A l'aisselle des premières feuilles, des bourgeons axillaires entrent alors en activité pour donner de nouvelles pousses : les talles. La 1^{ère} talle se forme à la base de la 1^{ère} feuille, la 2^{ème} talle à la base de la 2^{ème} feuille, et ainsi de suite.

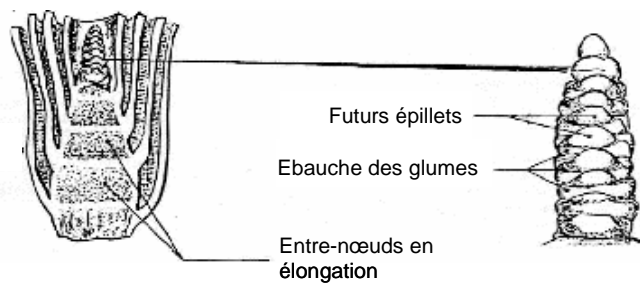


Figure 10 : Vue grossie d'un épi en formation, à la fin de l'initiation florale (d'après Soltner, 1988).

longueur du jour et de la date de semis. Divers paramètres abiotiques tels que les températures élevées, un stress hydrique ou un stress salin peuvent raccourcir ces phases de croissance du blé. Dans la suite de cette section, seuls les processus de la phase végétative importants pour le rendement et la phase reproductrice seront détaillés.

1.3.1 La phase végétative du développement du blé

La période végétative s'étend de la germination jusqu'au stade de montaison. La germination du grain de blé requiert un minimum de contenu en eau du grain de 35 à 45% du poids du grain. Elle peut avoir lieu de 4 à 37°C mais est optimale entre 12 et 25°C. La taille du grain n'affecte pas la germination. Cependant des gros grains permettent une croissance plus rapide des plantules, un nombre plus important de talles fertiles par plante et des rendements plus importants notamment en conditions de stress hydrique (Acevedo *et al*, 2002).

La mise en place et le développement des talles (ou tallage) est un processus en 2 étapes : l'initiation d'un méristème axillaire à l'aisselle des premières feuilles de la plantule de blé, sur le plateau de tallage (Figure 9), puis la croissance et l'élongation de la talle. Le nombre de talles potentielles varie avec le génotype (Soltner, 1988 ; Acevedo *et al*, 2002). Par exemple, les blés d'hiver et les variétés semi-naines produisent un plus grand nombre de talles. Toutes les talles ne produisent pas d'épis, beaucoup avortant avant l'anthèse. Le nombre de talles productives dépend du génotype et est fortement influencé par la densité de semis.

1.3.2 La phase reproductrice du développement du blé

1.3.2.1 L'initiation florale

Lorsque les plants de blé présentent de 4 à 8 feuilles, l'apex cesse de former des ébauches de feuilles et passe au stade reproductif. Cette étape s'appelle l'initiation florale. L'apex s'allonge et se segmente, formant les ébauches de futurs épillets (Figure 10). L'apex mesure alors 0,5 mm (Acevedo *et al*, 2002). Les primordia des glumes apparaissent à la base de chaque épillet (Soltner, 1988). Chaque épillet possède de huit à douze fleurs dans la partie centrale de chaque épi, et six à huit fleurs dans la partie distale et basale (Acevedo *et al*, 2002). Moins de la moitié de ces fleurs arriveront jusqu'à anthèse du fait d'avortement ou de développement incomplet des autres fleurs. A ce stade, le nombre d'épillets par épi est déjà déterminé. Ce stade est particulièrement sensible à des stress azoté et hydrique (Acevedo *et al*, 2002).



Figure 11 : Epi de blé tendre en floraison. La floraison, ou anthèse, s'observe à partir du moment où les étamines (jaunes) sortent des glumelles.

1.3.2.2 La montaison

La durée de cette phase est peu variable (de 28 à 30 jours). La croissance des talles s'arrête et les entre-nœuds s'allongent très rapidement. Les talles portant des épis en formation entrent alors en compétition avec ceux qui n'ont pu monter faute de nutriments ; ces derniers régressent et meurent. La croissance de l'épi débute dès l'apparition de l'avant-dernière feuille, et se poursuit jusqu'à 10 jours après anthèse. Le dernier nœud s'élève et la croissance de l'épi s'accélère dès que la ligule de la feuille drapeau commence à être visible. L'avortement des fleurs excédentaires a lieu à ce moment et serait probablement dû à une compétition pour le carbone à ce stade de développement (Acevedo *et al*, 2002). L'épi monte progressivement dans la gaine provoquant un gonflement de la gaine caractéristique. Cette phase s'arrête au moment de la différenciation des stigmates des fleurs et de la sortie de l'épi de la dernière feuille (épiaison).

1.3.2.3 L'épiaison

L'épiaison dure environ 32 jours. La formation des organes floraux s'achève avec la méiose qui donnera le pollen dans les anthères et le sac embryonnaire dans le carpelle. Chez le blé et l'orge, la méiose commence au milieu de l'épi et s'étend ensuite de part et d'autre de cette zone (Acevedo *et al*, 2002). Ce stade est très sensible aux stress environnementaux.

1.3.2.4 La floraison ou anthèse

La floraison ou anthèse s'observe à partir du moment où les étamines (et donc les anthères) sortent des glumelles de la fleur, et est prise comme point de référence pour la fécondation (Figure 11). Quand les anthères apparaissent, elles sont jaunes ; après exposition au soleil, elles deviennent blanches. L'anthèse commence dans la partie centrale de l'épi et s'étend ensuite de part et d'autre de cette zone. Cette période de floraison dure de trois à cinq jours. Dès lors, les fleurs de la zone centrale de l'épi sont fertilisées deux à quatre jours plus tôt que les fleurs en position distale, et les grains provenant des épillets centraux sont par conséquent généralement plus gros (Acevedo *et al*, 2002).

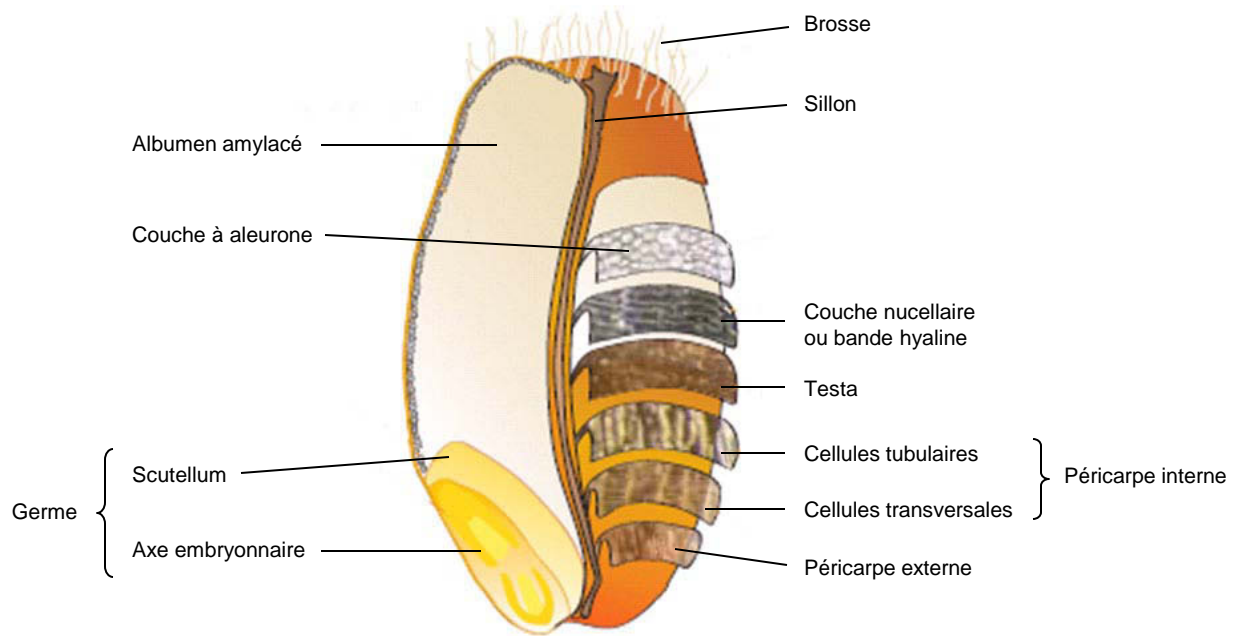


Figure 12 : Anatomie du grain de blé tendre (d'après Surget et Barron, 2005). Le grain de blé est constitué de trois parties : le germe (composé du scutellum et de l'axe embryonnaire), l'albumen (albumen amylicé et couche à aleurone) et les enveloppes (bande hyaline, testa, péricarpe interne et externe).

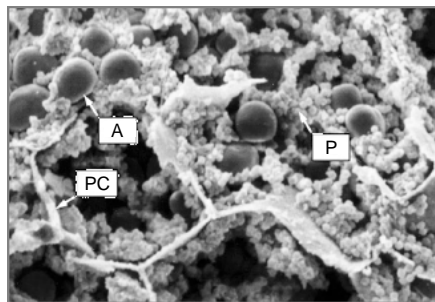


Figure 13 : Observation en microscopie électronique à balayage de l'albumen de maïs en développement (d'après Sabelli et Larkins, 2009). Les granules d'amidon (A) sont enchâssés dans une matrice protéique (P). PC : paroi cellulaire.

2 Le grain de blé

2.1 Structure du grain de blé

Les fruits de toutes les graminées sont des fruits secs indéhiscents appelés caryopses, dont les téguments de la graine sont soudés au péricarpe du fruit. Le caryopse de blé est nu puisque les glumelles n'adhèrent pas au grain. Le grain de blé est composé de trois parties : le germe, l'albumen et les enveloppes (Figure 12).

- Le germe est constitué d'une part de l'axe embryonnaire qui donnera la tigelle, le mésocotyle et la radicule de la future plantule de blé, et d'autre part du scutellum qui donnera le cotylédon (Evers et Millar, 2002 ; Surget et Barron, 2005). Le germe est la partie du grain où le taux d'humidité et la concentration en lipides sont les plus importantes (Pomeranz, 1988). A maturité, le germe compte pour 3% de la masse sèche du grain, avec approximativement des proportions égales de l'axe embryonnaire et du scutellum (Barron *et al*, 2007).
- L'albumen constitue le plus important compartiment du grain et représente environ 80% de son poids (Pomeranz, 1988). Il correspond au tissu de réserve. L'albumen à maturité est composé de deux parties: l'albumen amylacé et la couche à aleurone.

L'albumen amylacé est essentiellement constitué de granules d'amidon enchâssés dans une matrice protéique (Figure 13). Les cellules de l'albumen amylacé possèdent des parois fines et peuvent être classés en 3 grands groupes (Evers et Millar, 2002). De la partie périphérique jusqu'au centre de l'albumen, on trouve :

- les cellules périphériques, sous la couche à aleurone (environ 60 μm de diamètre)
- les cellules prismatiques, agencées en colonne (128 à 200 μm de long et 40 à 60 μm de large)
- les cellules centrales, de forme arrondie ou polygonale (72 à 144 μm de long et de 69 à 120 μm de large).

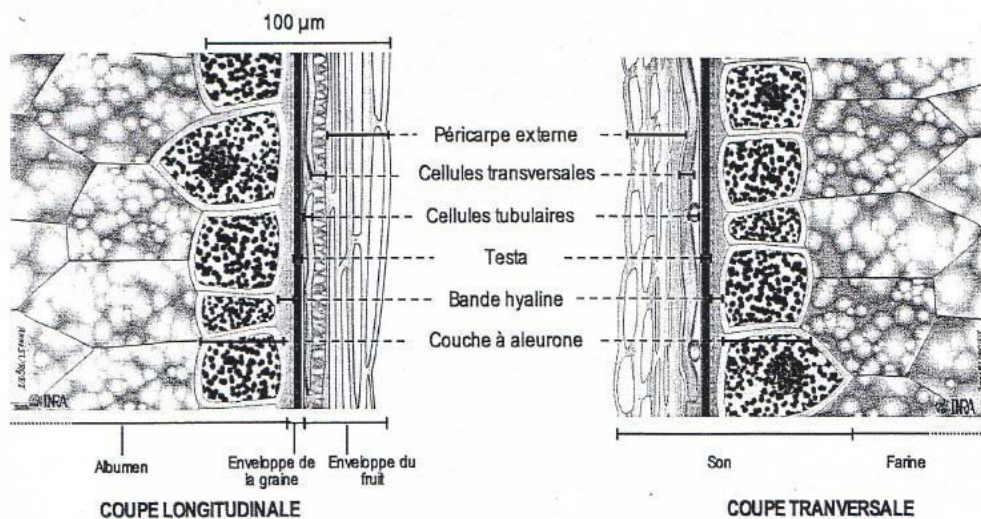


Figure 14 : Coupes longitudinale et transversale d'une section de grain de blé dans l'albumen et les enveloppes (d'après Surget et Barron, 2005).

L'albumen est la partie du grain de blé présentant le plus d'intérêt économique de part son utilisation finale. En effet, les protéines de réserves qui le constituent (gliadines et gluténines) ont la capacité de former en présence d'eau des liaisons covalentes et hydrogènes et des interactions notamment de type hydrophobe, aboutissant sous l'action du pétrissage à un réseau glutineux appelé gluten qui présente des propriétés viscoélastiques. De plus, l'amidon de l'albumen amylicé du grain de blé constitue une source d'énergie importante dans l'alimentation humaine et est également utilisé pour la production de biocarburant.

Une seule couche à aleurone entoure l'albumen amylicé chez le blé. Les cellules de la couche à aleurone sont de forme polygonale et mesurent approximativement 65 μm . Elles possèdent de gros noyaux et des parois épaisses (jusqu'à 8 μm). Les cellules de la couche à aleurone sont relativement riches en protéines (protéines de structure de type albumine et globuline), pentosanes (sucres à 5 carbones), vitamines (B1, B2, B3, B6, B9 et E), lipides et en minéraux (P, K, Mg, Mn et Fe) (Pomeranz, 1988 ; Antoine *et al*, 2002 ; McKeivith, 2004). Du fait de sa richesse en métabolites et de sa structure, la couche à aleurone assure un rôle nourricier et protecteur. Elle contient également beaucoup d'anthocyanes, responsables de la coloration du grain mature (Sabelli et Larkins, 2009). A maturité du grain, les cellules de la couche à aleurone sont les seules cellules vivantes restantes de l'albumen. Après l'imbibition de la graine, en réponse à une stimulation à l'acide gibbérellique par l'embryon, les cellules de la couche à aleurone vont exprimer des gènes d'enzymes protéolytiques et hydrolytiques, qui vont digérer les parois cellulaires de l'albumen amylicé et permettre la mobilisation de l'amidon et des protéines pour la croissance de l'embryon (Sabelli et Larkins, 2009).

- Les enveloppes sont composées de cinq tissus différents (Figure 14). On trouve de l'intérieur du grain vers l'extérieur : la couche nucellaire ou bande hyaline, la testa, les cellules tubulaires, les cellules transversales et le péricarpe externe. Les cellules tubulaires et les cellules transversales sont souvent également regroupées sous le nom de péricarpe interne.

La bande hyaline est une assise cellulaire entourant directement l'embryon et l'albumen. Les cellules ont une taille comprise entre 30 et 200 μm de longueur et entre 15 et 40 μm de largeur. Cette assise est tapissée d'une fine cuticule qui la relie à la

Tableau 1 : Distribution histologique des principaux constituants du grain de blé (d'après Feillet, 2000).

	Grain	Péricarpe (6)		Aleurone (7)		Albumen (84)		Germe (3)	
	%G	%T	%G	%T	%G	%T	%G	%T	%G
Protéines	13,7	10	4,4	30	15,3	12	73,5	31	6,8
Lipides	2,7	0	0	9	23,6	2	62,9	12	3,5
Amidon	68,9	0	0	0	0	82	100	0	0
Sucres réducteurs	2,4	0	0	0	0	1,8	62,7	30	37,3
Pentosanes	7,4	43	35,1	46	43,8	1,6	18,3	7	2,9
Cellulose	2,8	40	87,1	3	7,6	0,1	3,1	2	2,2
Minéraux	1,9	7	22,6	12	43,6	0,5	22,6	6	9,7

%G : % du constituant dans le grain ; %T : % du constituant dans le tissu. Les nombres entre parenthèses correspondent au % du tissu dans le grain.

L'amidon se retrouve exclusivement dans l'albumen amylacé ; les protéines sont retrouvées un peu partout mais leur teneur est particulièrement élevée dans le germe et la couche à aleurone ; les minéraux sont surtout présents dans la couche à aleurone, les pentosanes dans le péricarpe et la couche aleurone, les lipides dans le germe et la couche à aleurone ; et la cellulose est presque entièrement trouvée dans le péricarpe.

couche à aleurone. La bande hyaline semble jouer un rôle important dans la circulation de l'eau entre l'intérieur et l'extérieur du grain.

La testa est constituée de deux cuticules compressées, riches en lipides et composées de cellules allongées mesurant entre 120 et 190 μm de longueur et 20 μm de largeur (Bradbury *et al*, 1956) qui fusionnent par un film pigmentaire. Les axes des cellules de ces deux couches sont perpendiculaires entre eux ; l'un parallèle au sillon, l'autre perpendiculaire à celui-ci (Evers et Bechtel, 1988). La testa est très hydrophobe et permet au grain d'être peu perméable à l'eau.

Les cellules tubulaires et transversales du péricarpe interne sont respectivement parallèles et perpendiculaires à l'axe longitudinal du grain. Les cellules transversales sont propres aux graminées et sont de taille variable mesurant en moyenne 150 μm de longueur et 20 μm de largeur. Dans les grains de blé immatures, les cellules tubulaires et transversales possèdent des chloroplastes qui donnent la couleur verte aux grains immatures (Evers et Millar, 2002).

Le péricarpe externe, d'une épaisseur de 15 à 30 μm , est constitué de cellules mortes allongées mesurant de 80 à 300 μm de longueur.

Dans son ensemble, le péricarpe permet d'éviter les pertes en eau durant le développement du grain mais n'empêche pas sa pénétration (Evers *et al*, 1999). Les cellules mortes du péricarpe sont capables de retenir l'eau et d'augmenter le poids du grain de 4 à 5% après seulement quelques minutes d'imbibition (Hinton, 1955).

Contrairement au maïs et au riz, le grain de blé possède un sillon sur toute sa longueur et du côté du germe (Figure 12). Ce sillon résulte d'une invagination des téguments vers l'intérieur du grain. Les faisceaux nourriciers de la graine en cours de développement sont localisés au fond de ce sillon (Soltner, 1988).

La longueur d'un grain de blé est comprise entre 5 et 8 mm, sa largeur entre 2 et 4 mm ; son épaisseur entre 2,5 et 3,5 mm ; sa section longitudinale entre 10 et 16 mm^2 ; sa section transversale entre 4 et 7,5 mm^2 ; son poids entre 20 et 50 mg et sa densité entre 1,3 et 1,4. Néanmoins, la taille et le poids du grain dépendront également de sa position au sein de l'épi (Feillet, 2000 ; Calderini et Ortiz-Monasterio, 2003).

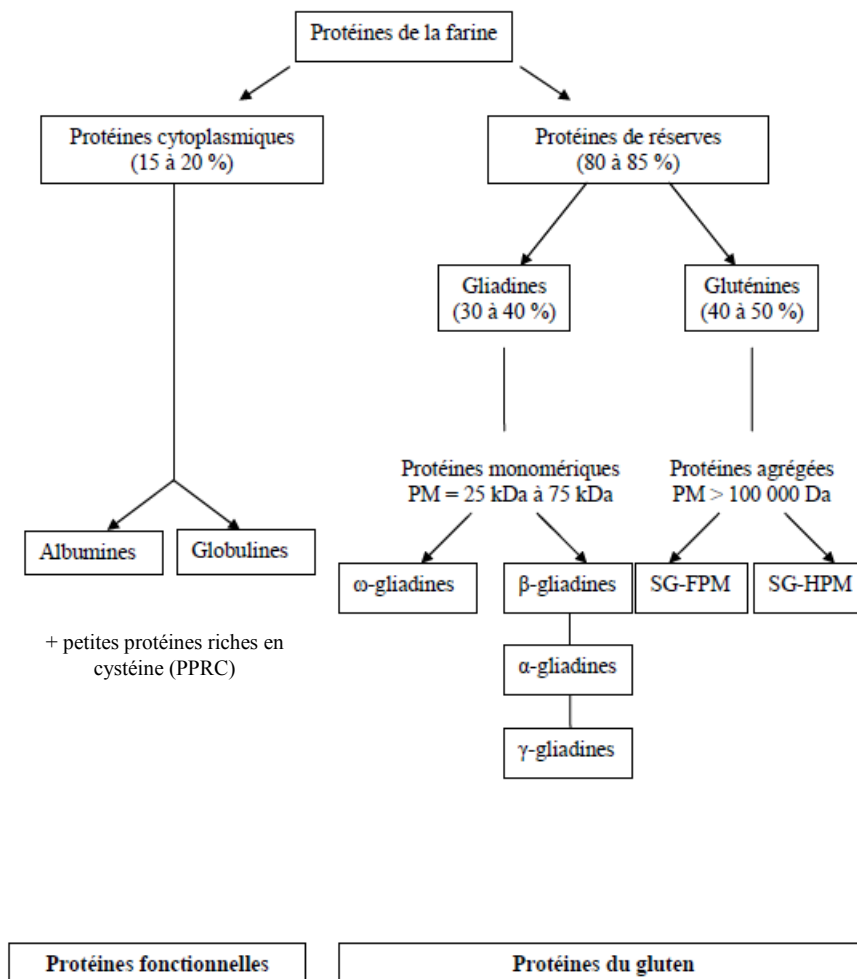


Figure 15 : Classification des protéines du grain de la farine du grain de blé : rapprochement entre les classifications d’Osborne et Shewry (d’après Feillet, 2000).

2.2 Composition du grain de blé

Les principaux composants du grain de blé sont, par ordre d'importance, l'amidon (70% de la matière sèche), les protéines (10 à 15% de la matière sèche), l'eau (12 à 14% de la matière fraîche) et les pentosanes (2 à 3% de la matière sèche). Les lipides, la cellulose, les sucres réducteurs, les minéraux et les vitamines sont minoritaires. Tous ces constituants se répartissent de manière inégale au sein des différentes fractions histologiques du grain (Tableau 1).

2.2.1 Les protéines du grain de blé

Les protéines du grain de blé sont généralement classées en fonction de leur solubilité (Osborne, 1924):

- albumines (solubles dans l'eau)
- globulines (solubles dans les solutions salines neutres)
- gliadines (solubles dans une solution d'alcool à 70%)
- gluténines (insolubles dans les solvants précités mais solubles dans des solutions acides ou alcalines, ou des détergents en présence d'un réducteur).

Cette classification a été revue en 1986 par Shewry et ses collaborateurs qui proposent les catégories suivantes (Figure 15) :

- les protéines métaboliques (albumines, globulines et les petites protéines riches en cystéine ou PPRC) ou cytoplasmiques
- les protéines de réserves (gliadines et gluténines).

Les albumines et les globulines représentent 15 à 20% des protéines de la farine. Elles sont constituées d'un grand nombre de protéines se différenciant par leur propriétés physico-chimiques et fonctionnelles (activités enzymatiques : α - et β -amylases, protéases, oxydo-réductases ; inhibiteurs d'enzymes, pouvoir émulsifiant et moussant, agent d'échange de liaisons disulfures) (Feillet, 2000). Les petites protéines riches en cystéine ou PPRC sont des petites protéines solubles dans un mélange chloroforme-méthanol (Feillet, 2000). Ce sont des protéines de transfert de lipides, des inhibiteurs d' α -amylases et de protéases, des lectines ou encore des protéines présentant des activités antifongiques et antibactériennes. Parmi les PPRC, on retrouve les puroindolines, protéines fixant les lipides, connues pour leur rôle dans la dureté du grain de blé.

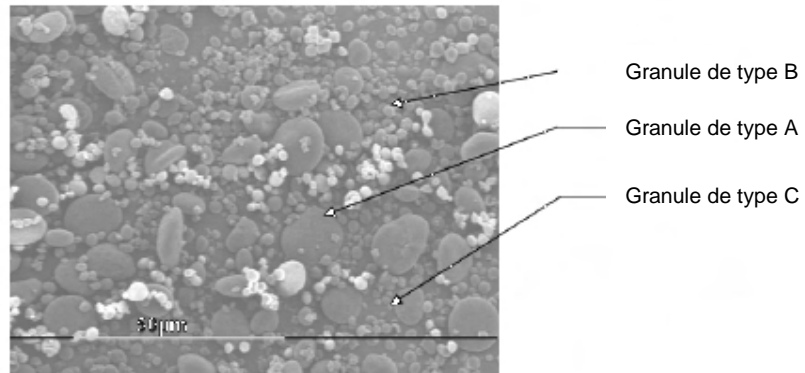


Figure 16 : Observation en microscopie électronique à balayage de granules d'amidon de blé tendre.

Les gliadines et gluténines sont également appelées prolamines car elles sont riches en l'acide aminé proline et représentent respectivement 30-40% et 40-50% des protéines de la farine. Elles sont les constituants principaux du gluten et constituent les protéines de réserves dans lesquelles la jeune plantule puisera les acides aminés dont elle a besoin au moment de la germination (Feillet, 2000). Les gliadines sont des protéines monomériques de 30 à 80 kDa et sont réparties en 4 classes : α -, β -, γ -, ω -gliadines. Les gluténines sont des protéines multimériques et les sous-unités les composant sont réparties en deux sous-groupes : les sous-unités gluténines de haut poids moléculaire (67-88 kDa) et les sous-unités gluténines de faible poids moléculaire (36-44 kDa). Il est admis que les gliadines sont responsables de la composante visqueuse du comportement rhéologique des pâtes et les gluténines de leur composante élastique (Feillet, 2000).

2.2.2 L'amidon

L'amidon est le principal constituant du grain de blé. Il est constitué de granules de type A, les plus gros (diamètre supérieur à 10 μ m) et majoritaires, de forme lenticulaire ; de granules de type B, plus petits (5-10 μ m) et sphériques ; et de granules de type C (taille inférieure à 5 μ m), sphériques (Bechtel *et al*, 1990 ; Raeker *et al*, 1998 ; Stoddard, 1999 ; Figure 16). Les cellules périphériques de l'albumen amylicé ne contiennent que des granules de type A tandis que les cellules prismatiques et centrales de l'albumen amylicé contiennent les trois types de granules (A, B et C). Les granules d'amidon de blé sont des très grands polysaccharides, constitués d'amylose et d'amylopectine, arrangés en structure semi-cristalline (Sabelli et Larkins, 2009). L'amidon est synthétisé à partir du saccharose, suite à sa conversion en ADP-glucose (Sabelli et Larkins, 2009). Il s'accumule dans les amyloplastes au cours de la maturation du grain et constitue la principale ressource énergétique dont dispose la plantule en début de germination. L'amidon est insoluble dans l'eau mais soluble dans le solvant di-méthylsulfoxyde (DMSO) (Feillet, 2000).

2.3 Développement du grain de blé

Les stades de développement du grain de blé sont estimés à partir du moment où la fécondation a eu lieu (soit l'anthèse) et peuvent être calculés de deux manières : en jours après anthèse (JAA) ou en degré-jours après anthèse ($^{\circ}$ CJ). En effet, le développement du grain est fortement influencé par la température (cf section 2.5) et ce dernier mode de calcul permet

Tableau 2 : Caractéristiques des différents stades de développement du grain de blé (d'après Rogers et Quatrano, 1983).

	Stades				
	I	II	III	IV	V
Jours après anthesis ^a	0-7	7-14	14-21	21-31	31-50
Couleur du grain	blanc	Blanc/ vert menthe	vert	vert	Brun
Longueur du grain (mm)	<5,5	5,5-6,5	6,5-7,0	6,5-7,0	6,5
Poids frais du grain	<20	20-35	35-50	50-60	40-50
Caractéristique de l'albumen	Aqueux	Laiteux, début de pâteux	Pâteux mou	Pâteux dur	Dur
Caractéristiques du péricarpe	Charnu	Contracté	Réduit et flexible	Sec et fin	Sec et fin
Longueur de l'embryon (mm) ^b	<0,4	0,4-1,5	1,5-2,5	2,5-2,7	2,5-2,7
a : <i>T. aestivum</i> cv Chinese spring cultivé avec 18h de lumière (19°C) et 6h d'obscurité (12°C)					
b : scutellum inclus					

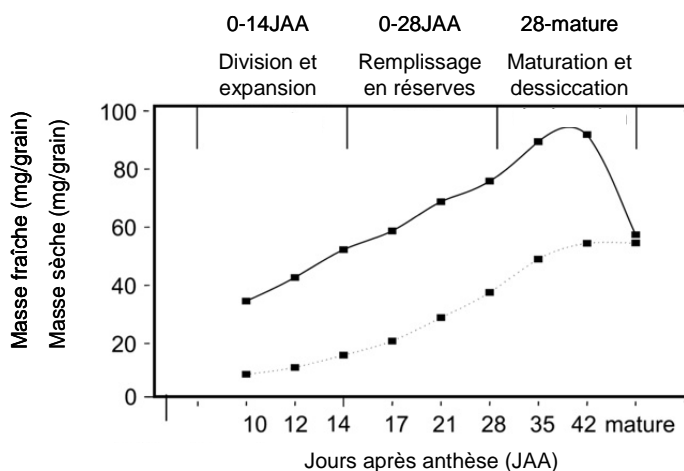


Figure 17 : Evolutions des masses fraîche et sèche de grains de blé tendre, cultivar Hereward (d'après Shewry et al, 2012).

une meilleure estimation du stade de développement puisqu'il tient compte des variations de température qu'il peut y avoir au cours du processus de développement. Les degré-jours après anthèse correspondent à la somme des températures moyennes journalières accumulées par le grain. Ce mode de calcul a été utilisé au cours de ce travail de thèse afin d'estimer les stades de développement du grain de blé.

Le développement du grain de blé est classiquement subdivisé en trois grandes phases, sur la base du développement de l'albumen : la phase de division cellulaire, la phase de remplissage en réserves, et la phase de maturation-dessiccation du grain. Rogers et Quatrano (1983) ont par la suite proposé une subdivision du développement du grain de blé en 5 phases (Tableau 2) : la phase précoce de la division ou phase I (de 0 à 7 JAA ce qui correspond environ à un intervalle de 0 à 140°CJ pour une température moyenne de 20°C), la phase tardive de la division ou phase II (de 7 à 14 JAA ; 140 à 260°CJ), la phase précoce du remplissage ou phase III (de 14 à 21 JAA ; 260 à 400°CJ), la phase tardive du remplissage ou phase IV (de 21 à 31 JAA ; 400 à 600°CJ) et la phase de maturation ou phase V (de 31 à 50 JAA ; 600 à 1000°CJ).

L'albumen en développement va subir divers processus biologiques : la double-fécondation, la formation d'un syncytium, la cellularisation, la différenciation cellulaire, la division cellulaire, l'endoréduplication, l'accumulation des réserves et pour finir la maturation du tissu comprenant la mort cellulaire programmée, la dormance et la dessiccation (Sabelli et Larkins, 2009). Ces processus biologiques vont être mis en place graduellement et ne s'arrêtent pas automatiquement à la transition de deux phases. Dès lors, les limites chronologiques entre les différentes phases sont parfois floues et peuvent varier selon la littérature. De plus ces divers processus vont se chevaucher dans le temps.

Au cours de la phase de division cellulaire (0-14 JAA), la structure basique du grain est établie et le processus de remplissage du grain est initié. La longueur du grain est fixée à la fin de cette phase (Rogers et Quatrano, 1983). De 14 à 28 JAA, l'amidon et les protéines s'accumulent fortement, et la masse sèche du grain augmente de près de 2 fois. Le remplissage en réserves ralentit après 28 JAA et est achevé à 35 JAA. Les dimensions largeur et épaisseur du grain atteignent leur maximum à la fin de cette phase de remplissage. La masse fraîche du grain chute rapidement après 42 JAA, suite au processus de dessiccation (Shewry *et al*, 2012 ; Figure 17).

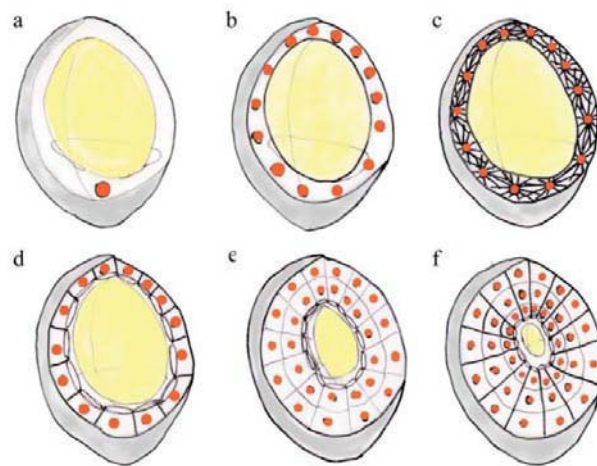


Figure 18 : Représentation schématique du processus de cellularisation de l'albumen (d'après Olsen, 2001). (a) Le noyau triploïde fertilisé (en orange) vient se placer à l'extrémité proximale du cytoplasme de la cellule centrale, puis subit des divisions successives sans cytokinèse (b). (c) La cellularisation de l'albumen va être initiée par la formation de réseaux microtubulaires radiaux (RMR) entre les noyaux. (d) La formation des parois cellulaires est facilitée par l'interaction entre les RMR, et chaque cellule se retrouve alors entourée d'une structure pariétale en tube, ou encore appelée « en alvéole ». (e-f) Des divisions périclines vont avoir lieu jusqu'à fermeture complète de la vacuole centrale (en jaune).

2.3.1 *La phase de division*

Le grain résulte de la double-fécondation du sac embryonnaire. Un des deux noyaux du grain de pollen va féconder l'oosphère pour donner le zygote diploïde, et le deuxième noyau va féconder les deux noyaux polaires de la cellule centrale du sac embryonnaire pour donner le tissu nourricier triploïde, l'albumen. Des divisions successives de la cellule triploïde sans cytokinèse pendant 72 heures aboutissent à la formation d'un coenocyte à plus de 2000 noyaux (Chojewski *et al*, 1986). Les noyaux se répartissent à la périphérie du coenocyte, autour d'une large vacuole centrale. L'albumen coenocytique va alors subir un processus de cellularisation pendant 24 heures. Des réseaux microtubulaires radiaux vont se former entre les noyaux, et les parois cellulaires vont se créer sur ces réseaux (Olsen, 2001 ; Figure 18).

Au cours des 10 jours suivants, des processus de différenciation, division et expansion cellulaires vont avoir lieu pour aboutir à un albumen possédant jusqu'à 300 000 cellules. Des divisions périclines vont permettre la formation centripète de plusieurs assises cellulaires, de structure alvéolaire caractéristique de ce stade, jusqu'à la disparition de la vacuole centrale (Olsen, 2001 ; Sabelli et Larkins, 2009). La différenciation cellulaire va permettre la formation des types cellulaires principaux de l'albumen en développement : les cellules de transfert, les cellules de la couche à aleurone, les cellules de l'albumen amylicé et les cellules de la région entourant l'embryon (ou cellules ESR, Embryo-surrounding region cells) (Sabelli et Larkins, 2009). Les cellules de transfert présentent un cytoplasme dense, riche en mitochondries sphériques de petite taille. Elles sont localisées au-dessus de la projection nucellaire et facilitent l'accumulation rapide de nutriments (principalement saccharose et acides aminés) par l'albumen (Sabelli et Larkins, 2009). Les cellules ESR constituent plusieurs couches de cellules entourant l'embryon. Elles se différencient dès la fin de la cellularisation. Ces cellules sont supposées être mitotiquement très actives et impliquées dans l'apport de sucres à l'embryon, par une voie apoplastique (Cossegal *et al*, 2007). Ces cellules seraient également impliquées dans la défense contre des pathogènes et dans la signalisation à l'interface embryon-albumen. Au fur et à mesure que l'embryon grandit, les cellules ESR vont progressivement disparaître, pour ne rester qu'à l'état de vestiges à la base de l'albumen à 12 JAA. Une seule couche de cellules à aleurone va être différenciée à la périphérie de l'albumen du grain de blé. Dans les cellules de l'albumen amylicé, la formation des granules d'amidon de type A va commencer dès la fin de la cellularisation, vers 3-4 JAA, et s'arrêter brusquement à 7 JAA. Une phase de division cellulaire massive va avoir lieu, de la partie centrale de l'albumen vers la périphérie, jusqu'à 8-12 JAA dans la partie centrale de l'albumen, et jusqu'à 20 à 25 JAA dans les cellules de la couche à aleurone. L'index

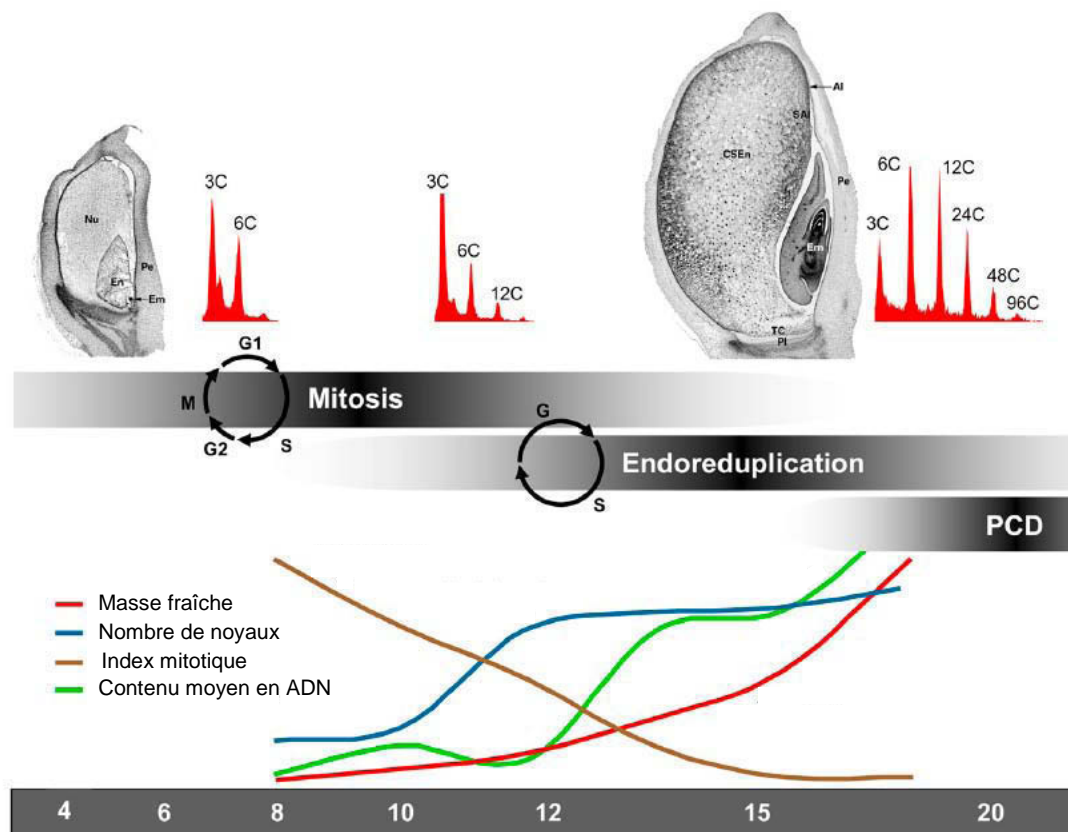
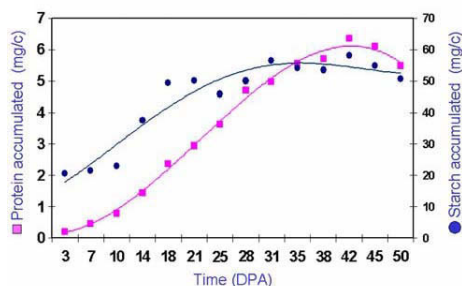


Figure 19 : Développement de l'albumen de maïs de 4 à 20 JAA (d'après Sabelli et Larkins, 2009). Bien que cette figure traite du développement de l'albumen chez le maïs, les auteurs suggèrent qu'elle est un bon exemple pour le développement de l'albumen d'autres céréales, comme le blé tendre. De 4 à 20 JAA, l'albumen va subir une phase de division cellulaire massive, suivie d'une phase d'endoréduplication, puis d'une phase de mort cellulaire programmée (PCD). La phase d'endoréduplication coïncide avec la croissance importante et rapide de l'albumen et l'accumulation en réserves. Le contenu en ADN de l'albumen à trois moments du cycle de développement est représenté par les profils de cytométrie de flux en rouge.

A



B

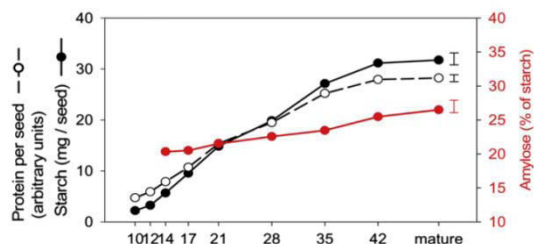


Figure 20 : Accumulation des protéines et de l'amidon au cours du développement de l'albumen, dans le blé tendre cultivars Bobwhite (A) et Hereward (B) (d'après Laudencia-Chinguanco *et al*, 2007 (A) et Shewry *et al*, 2012 (B)). (A) L'accumulation des protéines et de l'amidon figurent en bleu et rose, respectivement. (B) L'accumulation des protéines et de l'amidon sont représentés en noir en pointillés et en trait plein, respectivement ; l'accumulation de l'amylose figure en rouge.

mitotique culmine autour de 8 à 10 JAA et décline ensuite brusquement (Sabelli et Larkins, 2009).

Autour de 10 JAA, les cellules de l'albumen vont basculer de la mitose à des cycles d'endoréduplication, au cours desquels des cycles réitérés de synthèse d'ADN vont avoir lieu sans condensation de la chromatine, ni ségrégation des chromatides sœurs et sans cytokinèse ; conduisant à la formation de noyaux polyploïdes (Sabelli, 2012). Tout comme la division cellulaire, la phase d'endoréduplication va avoir lieu du centre de l'albumen vers la périphérie (Sabelli, 2012). L'augmentation du contenu en ADN des cellules, suite à l'endoréduplication, est corrélée à la taille des noyaux et des cellules, à la croissance rapide du caryopse, ainsi qu'à la synthèse et à l'accumulation des réserves (Sabelli et Larkins, 2009 ; Figure 19). D'après la littérature, l'endoréduplication dans l'albumen permettrait d'augmenter la quantité de matrice pour supporter des taux de transcription forts, ou encore permettrait une expansion cellulaire et une croissance tissulaire rapide sans division cellulaire (Sabelli et Larkins, 2009). Une modulation de l'activité des CDK (cyclin-dependent kinases) apparaît notamment importante pour cette transition (Sabelli, 2012).

La transition de la phase de division à la phase de remplissage est accompagnée d'une importante reprogrammation des profils d'expression des gènes (Sreenivasulu *et al*, 2004 ; Drea *et al*, 2005 ; Laudencia-Chingcuanco *et al*, 2007 ; Wan *et al*, 2008 ; Capron *et al*, 2012) et semble notamment régulée en partie par le saccharose et l'activité de la sucrose synthase (SuSy).

2.3.2 La phase de remplissage

Les protéines de réserves et l'amidon vont s'accumuler selon un profil similaire dans le temps (Shewry *et al*, 2012 ; Figure 20). Le dépôt de protéines de réserves dans l'albumen va commencer de manière sporadique à 10 JAA, puis augmenter fortement de 14 à 21 JAA et enfin ralentir jusqu'à 42 JAA. Les cellules sous la couche à aleurone vont accumuler beaucoup plus de protéines que les cellules centrales de l'albumen, certaines d'entre-elles possédant la totalité de leur contenu protéique à 22 JAA (Shewry *et al*, 2012 ; Figure 21). La fraction protéique de ces cellules est riche en gliadines et gluténines de faible poids moléculaires (Tosi *et al*, 2009, 2011). Au même moment, tout au long de la phase de remplissage, la proportion des protéines métaboliques albumines et de globulines va décliner (Hurkman *et al*, 2009). Le contenu en amidon va augmenter rapidement de 14 à 35 JAA et ralentir jusqu'à 42 JAA. Contrairement au profil spatial d'accumulation protéique, les cellules

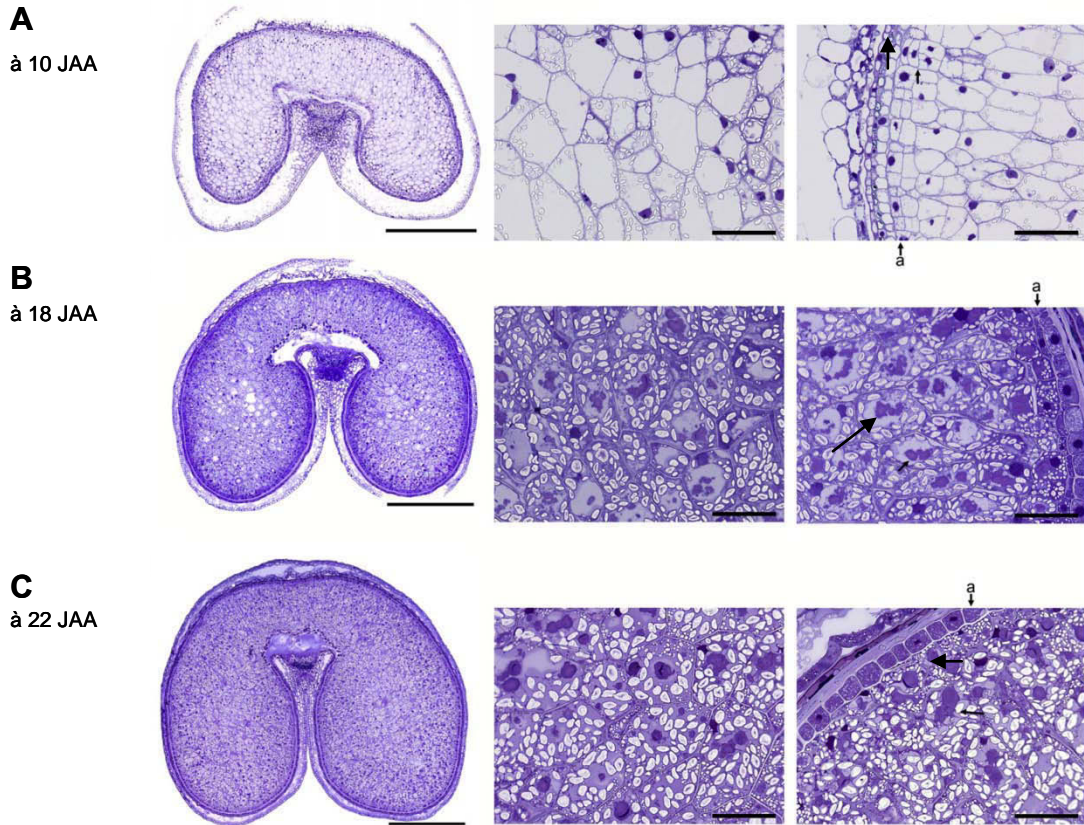


Figure 21 : Sections transversales de grain de blé tendre en développement, cultivar Hereward, avant et pendant la phase de remplissage en réserves (d'après Shewry *et al*, 2012). A gauche sont représentés des sections transversales entières de grain ; et au milieu et à droite, des images agrandies de la partie centrale et périphérique de l'albumen. La couche à aleurone est indiquée par la lettre « a ». **(A)** La flèche indique des cellules sous la couche à aleurone qui se sont divisées récemment. **(B)** La flèche indique des dépôts protéiques dans la vacuole d'une cellule sous la couche à aleurone. **(C)** La flèche indique une cellule sous la couche à aleurone entièrement remplie de protéines. Les sections sont colorées au bleu de toluidine. Barre = 1mm dans le panneau gauche ; 100 nm dans les panneaux centraux et droite.

sous la couche à aleurone accumuleront moins d'amidon que les cellules centrales de l'albumen (Shewry *et al*, 2012). Des granules de type A vont d'abord être formés, puis des granules de type B et finalement C (Briarty *et al*, 1979 ; Bechtel *et al*, 1990).

Le ratio acide abscissique/auxine (ABA/Aux) influence le remplissage en réserves de l'albumen (Yang *et al*, 2001a ; Xu *et al*, 2007). Le gène *SnRK1* (sucrose non-fermenting-1 related protein kinase 1), un gène codant une protéine kinase homologue de SNF1 (Sucrose non-fermenting 1) chez la levure et AMPK (AMP-activated protein kinase) chez les mammifères, est de plus en plus considéré comme un bon candidat pour l'intégration des signalisations sucre et acide abscissique et la mise en place de la synthèse d'amidon dans l'albumen (Sabelli et Larkins, 2009). SnRK1 serait induit en réponse au saccharose et pourrait affecter la synthèse d'amidon en régulant à la fois l'expression du gène *SuSy* et l'activité de l'AGPase (Halford et Paul, 2003). La sous-famille SnRK1b est particulièrement intéressante car elle est spécifique des céréales et fortement exprimée dans le grain. Chez le riz, le sorgho et le maïs, l'expression de *SnRK1b* est associée avec le développement de la capacité puits et la région des cellules de transfert dans l'albumen (Kanegae *et al*, 2005 ; Jain *et al*, 2008) et semble jouer un rôle important en faisant la liaison entre la signalisation ABA/Saccharose et la transition vers la phase de remplissage (Sreenivasulu *et al*, 2006), et/ou en inhibant la division cellulaire.

Très récemment, une analyse en composantes principales des profils d'accumulation de métabolites solubles, dans le cultivar Hereward, a permis de montrer que les grains accumulent également des métabolites solubles (comme par exemple les sucres simples ou les acides aminés) selon un procédé en 2 étapes : une première étape de 14 à 35 JAA, i.e durant toute la phase de remplissage et le début de la phase de maturation ; et une deuxième étape de 35 à 42 JAA, i.e à la phase de maturation (Shewry *et al*, 2012). Le taux de fructose diminue brusquement de 14 à 35 JAA et les taux de glucose et de maltose augmentent de 14 à 35 JAA puis chutent brutalement après 42 JAA. La bétaine est connue pour agir comme un osmorégulant chez de nombreux organismes dont les plantes. Les auteurs suggèrent que la forte accumulation de ce métabolite à la fin de la phase de maturation doit sans doute jouer un rôle osmoprotecteur dans le grain de blé.

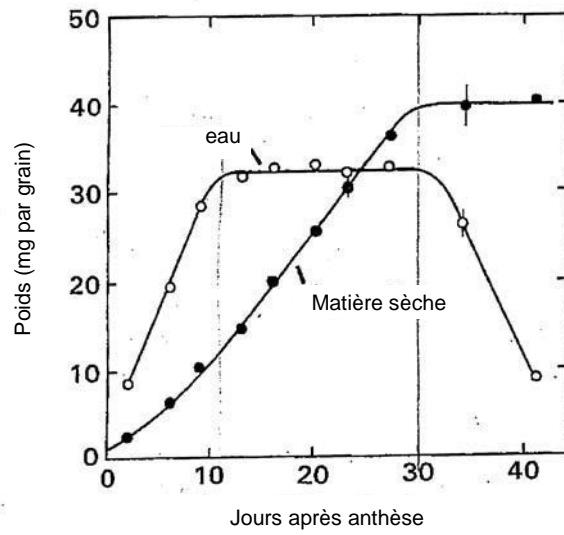


Figure 22 : Dynamique de la matière sèche et du contenu en eau des grains basaux d'épillets centraux de blé tendre, cultivar Achill (d'après Schnyder et Baum, 1992).

2.3.3 La phase de maturation

La phase de maturation entraîne l'arrêt de voies métaboliques et comprend les processus biologiques suivants : la mort cellulaire programmée, la dormance et la dessiccation.

La mort cellulaire programmée de l'albumen commence lentement vers 16 JAA et culmine vers 30 JAA (Young et Gallie, 2000 ; Sabelli, 2012). Contrairement au maïs, le programme de mort cellulaire de l'albumen chez le blé n'est pas initié à un endroit précis du tissu mais plutôt de manière stochastique (Young et Gallie, 2000). Toutes les cellules de l'albumen, excepté les cellules de la couche à aleurone, vont subir une fragmentation de l'ADN, une condensation de la chromatine et un désassemblage de la membrane nucléaire. La mort cellulaire programmée faciliterait l'hydrolyse des nutriments pour leur absorption future par l'embryon au cours de la germination (Nguyen *et al*, 2007). La mise en place et la progression du programme de mort cellulaire de l'albumen ont été décrites chez le maïs comme étant influencées par la signalisation à l'éthylène et l'ABA (Sabelli et Larkins, 2009).

La dormance est une inhibition du programme de germination de la graine, afin que la germination puisse se faire lorsque les conditions environnementales sont favorables. L'ABA est également impliqué dans ce processus de mis en latence du grain (Sabelli et Larkins, 2009) et des études ont montré l'importance des gènes de la classe viviparous dans la dormance ; montrant notamment que l'expression de l' α -amylase était réprimée par le gène *VIVIPAROUS1* (Hoecker *et al*, 2009).

Au cours de la dessiccation, le grain va perdre très rapidement son eau ; et celle-ci va être substituée au sein des cellules par l'amidon et les protéines (Berger, 1999). Le grain va alors atteindre sa masse sèche maximale. Cet état est appelé la maturité physiologique du grain de blé.

2.4 Le statut hydrique du grain au cours de son développement

Le contenu en eau du grain augmente fortement dans les premiers JAA et atteint son maximum vers 10-14 JAA. Il se stabilise alors pendant 2 à 3 semaines (Schnyder et Baum, 1992 ; Figure 22), c'est le plateau hydrique. Le contenu en eau chute brutalement à 30 JAA.

Le statut hydrique du grain est intimement lié au développement de celui-ci tout au long des 3 grandes phases. La croissance initiale des grains, dans les phases précoces du développement, est en grande partie due à l'accumulation rapide d'eau dans le grain

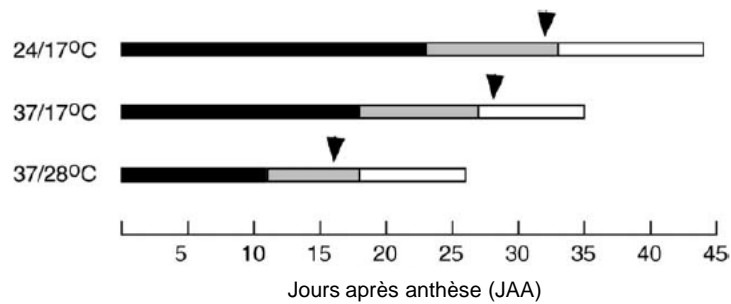


Figure 23 : Durées de développement de grain de blé de plants cultivés sous trois conditions de température, de l'anthèse jusqu'à maturité (d'après Dupont et Altenbach, 2003). Les températures moyennes diurnes et nocturnes sont indiquées respectivement à gauche du graphique. Les barres noires, grises et blanches représentent respectivement le temps nécessaire pour atteindre le contenu maximal en eau du grain, la masse sèche maximale du grain et la maturité du grain, respectivement. La flèche noire représente le déclenchement de l'apoptose.

(Schnyder et Baum, 1992). De plus, l'arrêt de la division cellulaire a lieu a peu près au même moment que l'arrêt de cette accumulation rapide d'eau dans le grain (Nadaud *et al*, 2010). En outre, la taille finale des cellules serait influencée en partie par l'accumulation en eau (Egli, 1998). Schnyder et Baum (1992) ont proposé que le contenu en eau détermine le volume de croissance du grain au cours de la phase de division cellulaire de l'albumen. Au cours de la phase de remplissage, le statut hydrique et la dynamique d'accumulation de la matière sèche sont étroitement liés (Schnyder et Baum, 1992 ; Calderini et Reynolds, 2000). La fin d'accumulation d'eau dans l'albumen coïncide notamment avec la fin de l'accumulation de la matière sèche (Schnyder et Baum, 1992). Au cours de la dessiccation du grain, le statut hydrique du grain chute à 10-15%. La réduction des activités métaboliques au cours de la phase de maturation semble notamment causée par cette perte massive d'eau (Young et Gallie, 2000).

2.5 L'effet des stress abiotiques sur le développement du grain

Les stress abiotiques regroupent toutes les conditions environnementales et leurs combinaisons qui affectent négativement l'expression du potentiel génétique de croissance, de développement et de reproduction (Jones et Qualset, 1984).

De nombreuses études rapportent les effets de l'environnement sur le développement du grain (revues de Dupont et Altenbach, 2003 ; Altenbach, 2012 ; Thitisaksakul *et al*, 2012).

La température altère la durée et la vitesse du remplissage du grain. Au-delà d'une température moyenne journalière de 20°C, la durée du remplissage du grain est réduite (Wiegand et Cuellar, 1981). On estime d'ailleurs que le poids du grain est réduit de 4 à 8% par °C supérieur à la gamme de température de croissance optimale (12 à 26°C) (Acevedo *et al*, 2002). La température peut également plus ou moins impacter la vitesse de remplissage en réserves en fonction de la variété considérée (Sofield *et al*, 1977a ; Tashiro et Wardlaw, 1990 ; Wardlaw et Moncur, 1995). Pour certaines variétés, la vitesse de remplissage en réserves augmente avec la température jusqu'au seuil de 30°C, tandis que pour d'autres variétés, la vitesse de remplissage reste constante quelque soit la température. Altenbach et ses collaborateurs (2003) ont montré que des fortes températures impactaient le minutage de tous les stades du développement du grain ; le programme de développement en entier étant accéléré et compressé (Figure 23). Dans des conditions de fortes températures, l'accumulation de l'amidon et des protéines commence plus tôt et le temps pour atteindre le contenu maximal

en eau du grain et les poids frais et sec maximaux sont raccourcis. L'apoptose est également déclenchée plus tôt (Dupont et Altenbach, 2003 ; Figure 23). Les fortes températures réduisent également le nombre final de cellules de l'albumen (Acevedo *et al*, 2002).

Plusieurs études montrent que le stress hydrique diminue la taille du grain, en raccourcissant la période de remplissage du grain mais pas la vitesse d'accumulation des réserves (Brooks *et al*, 1982 ; Nicolas *et al*, 1984 ; Kobata *et al*, 1992 ; Altenbach *et al*, 2003).

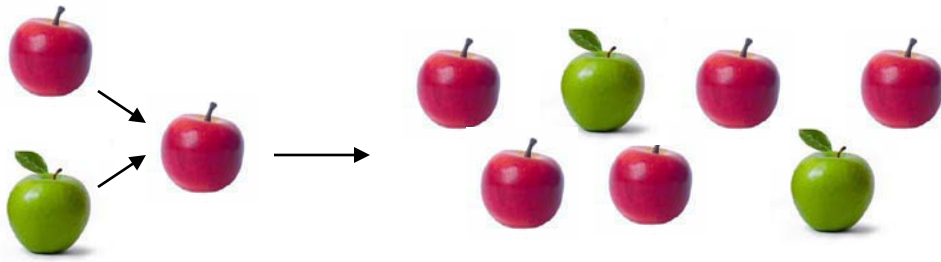
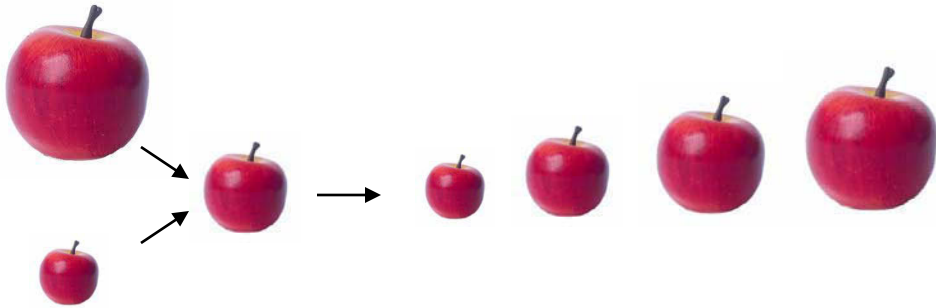
A**B**

Figure 24 : Notions de caractères simple (A) et complexe (B) (d'après Caromel B, INRA). (A) Dans le cas d'un caractère simple (caractère qualitatif, mendélien) comme la couleur de la pomme dans cet exemple, les diverses variantes de ce caractère dans une population se distinguent nettement les unes des autres (variation discontinue) et sont en nombre réduit. L'expression des caractères qualitatifs dépend principalement de gènes dits majeurs sur chacun desquels l'influence de l'environnement est nulle ou négligeable. (B) Dans le cas d'un caractère complexe (caractère quantitatif) comme la taille de la pomme dans cet exemple, les diverses variantes de ce caractère dans une population sont difficilement classifiables (variation continue) et sont en très grand nombre, au point que chaque variante ne décrit qu'un faible nombre d'individus. Un caractère quantitatif dépend généralement de plusieurs gènes et est très affecté par le milieu. Dans le cas de ces caractères, il existe des gènes majeurs qui contrôlent l'expression de ces caractères et des gènes mineurs qui sont supposés très nombreux pour chaque caractère quantitatif et qui auraient chacun un effet faible mais additif.

3 Le rendement en grains

Le rendement en grains chez les céréales (exprimé en unité de masse.hectare⁻¹ à 0% d'humidité) est un caractère quantitatif complexe, d'intérêt agronomique important. Pour de nombreuses raisons (changements climatiques, augmentation croissante de la population mondiale, demande accrue pour la production de biocarburants), le rendement en grains est actuellement la principale cible des programmes d'amélioration variétale (Sreenivasulu et Schnurbusch, 2012).

Contrairement aux caractères qualitatifs qui présentent une variation discrète (ou discontinue) du phénotype, les caractères quantitatifs, comme le rendement, présentent une variation phénotypique continue dans les populations, illustrée par une distribution gaussienne de la valeur du caractère (De Vienne, 1995). Pour de tels caractères, les individus ne sont pas classés mais mesurés (Figure 24). En opposition aux caractères qualitatifs qui sont le plus souvent à déterminisme monogénique, les caractères quantitatifs sont à déterminisme multigénique. La distribution continue d'un caractère quantitatif s'explique par l'effet de plusieurs loci en ségrégation sur la variation du caractère, auquel s'ajoutent le plus souvent des effets de l'environnement. Dans le contrôle de ce type de caractère, les gènes ne présentent pas classiquement des profils de dominance et récessivité, mais la contribution de chacun au sein du locus serait plutôt additive.

Ainsi, le rendement en grains est contrôlé par ces locis à effets quantitatifs, plus communément appelés QTL (Quantitative Trait Loci), dérivés de variations naturelles. Pour définition, un QTL est une région chromosomique pouvant comporter plusieurs gènes contrôlant un caractère quantitatif. D'une manière générale, la recherche de QTL est basée sur l'utilisation de cartes génétiques. Elle consiste en la recherche de liaison statistique entre des allèles aux marqueurs et le phénotype des plantes pour le caractère. Cependant, l'hérédité polygénique du rendement en grains inclut des interactions fortes avec l'environnement (le rendement est un caractère dit à hérédité multifactorielle) ; et les interactions génotypes x environnement peuvent fortement affecter la détection des QTLs de rendement. Ainsi l'ampleur des effets des QTLs peut différer d'un milieu à l'autre ; et des QTLs détectés dans un milieu peuvent ne plus l'être dans un autre. Cet aspect est appelé la stabilité des QTLs : des QTLs retrouvés dans tous ou presque tous les environnements testés sont appelés QTLs

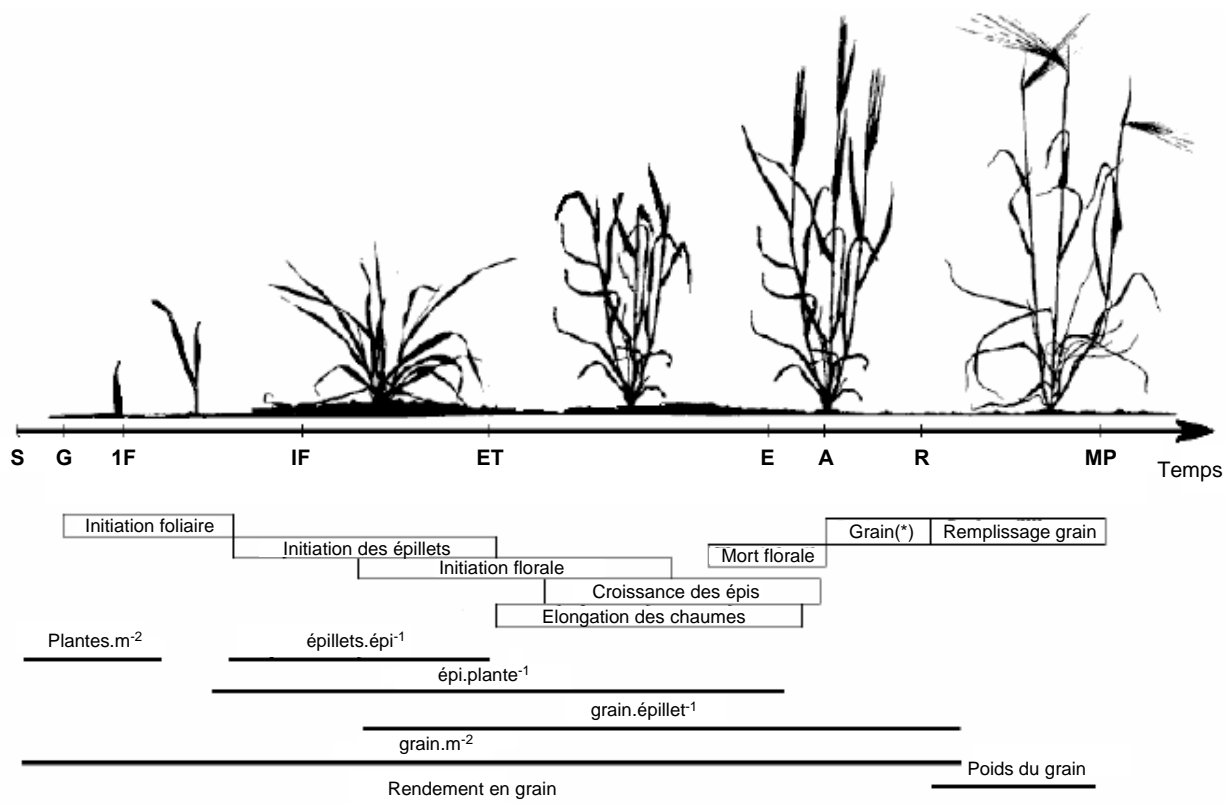


Figure 25 : Schéma des phases de croissance et de développement du blé et périodes d'établissement des différentes composantes du rendement en grain (d'après Acevedo *et al*, 2002). S, semis ; G, germination ; 1F, stade 1 feuille ; IF, initiation florale ; ET, épillet terminal formé sur le méristème inflorescentiel ; E, épiaison ; A, anthèse ; R, début du remplissage du grain ; MP, maturité physiologique. (*) : Etablissement de la structure basique du grain. Les rectangles représentent les périodes de différents processus de développement de certains organes ; les périodes déterminant les différentes composantes du rendement en grain sont représentées par des traits pleins en dessous de la frise chronologique.

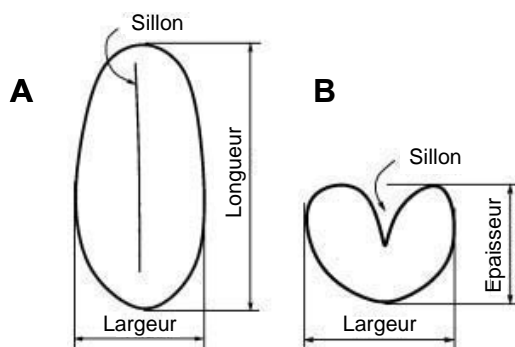


Figure 26 : Dimensions du grain de blé: longueur, largeur et épaisseur (d'après Sun *et al*, 2007). (A) vue de dessus, (B) section transversale.

‘généralistes’ ou majeurs, tandis que d’autres, les QTLs ‘spécialistes’ ou mineurs, sont spécifiques d’un seul environnement (De Vienne, 1995).

Chez le blé, le rendement en grains est la résultante de plusieurs composantes, elles-mêmes sous le contrôle de QTL, mises en place tout au long des stades végétatifs et reproductifs de la plante cultivée (Figure 25). Ces caractéristiques mesurables sont entre autre :

- le nombre de plantes par m²
- le nombre de talles par plante
- le nombre d’épis par plante
- le nombre d’épillets par plante
- le nombre de grains par épi
- le poids de 1000-grains (PMG)

Le nombre de grains par plante et le poids du grain (ou poids de 1000-grains, PMG) sont les composantes les plus importantes du rendement. Le nombre final de grains par plante est déterminé par le nombre d’épis fertiles par plante, le nombre d’épillets fertiles par épi et le nombre de grains par épillet (Sreenivasulu et Schnurbusch, 2012). La composante poids du grain est largement déterminée par la taille du grain et son degré de remplissage. La taille du grain de blé est spécifiée par ses trois dimensions : la largeur, la longueur et l’épaisseur du grain (Campbell *et al*, 1999 ; Dholakia *et al*, 2003 ; Figure 26). Ces trois composantes sont toutes positivement corrélées au poids de 1000-grains. Cependant, la largeur du grain reste la composante la plus importante du PMG (Sun *et al*, 2009 ; Gegas *et al*, 2010). La largeur du grain n’est pas exclusivement reliée qu’au rendement en grains mais a été reportée comme présentant également des effets sur la vigueur et la précocité de croissance de la plantule de blé, permettant ainsi une stabilisation de la capacité de production.

L’étude des déterminants génétiques du rendement chez le blé est complexe du fait notamment de la nature intégrative de ce caractère. En effet, il existe de nombreuses relations entre les différentes composantes. Il existe par exemple une relation négative entre le nombre de grains par unité de surface et la masse individuelle des grains (Fisher, 2007). De même, d’une manière générale chez les céréales, le nombre de talles est négativement corrélé à la performance de rendement en grains (Khush, 2001). Très fréquemment dans les études, le déterminisme génétique du rendement est analysé en décomposant et étudiant séparément chacune des composantes.

Tableau 3 : Exemples de QTLs détectés pour différentes composantes du rendement, chez le blé tendre.

Composante	Nombre de QTLs détectés, nom du QTL	Chromosome(s)	Référence bibliographique
Rendement en grain	4 QTL	5A	Kato <i>et al</i> , 2000
Nombre de talles.plante ⁻¹	3 QTL	5A	
Poids total en grain sur l'épi	4 QTL	5A	
Poids de 1000-grain	2 QTL	5A	
Nombre d'épillets.épi ⁻¹	4 QTL	5A	
Hauteur de la plante	6 QTL	2D, 4B, 4D, 5B, 7A, 7B	McCartney <i>et al</i> , 2005
Temps pour atteindre la maturité complète	4 QTL	3B, 4A, 4D, 7D	
Résistance à la verse	3 QTL	3D, 4B, 4D	
Rendement en grain	5 QTL	2A, 2B, 3D, 4A, 4D	
Poids de 1000-grain	6 QTL	2A, 3D, 4A, 4B, 4D, 6D	
Rendement en grain	1 QTL, <i>Qgyld.agt-1B</i>	1B	Kuchel <i>et al</i> , 2007
Largeur du grain	1 QTL	1B	Bresseghello et Sorrells, 2007
Volume du grain	7 QTL	1B, 2B, 2D, 3A, 4A, 5B, 7B	
Poids du grain	3 QTL	2B, 2D, 4A	
Longueur du grain	3 QTL	4B, 5B, 5D	
Poids du grain	1 QTL, <i>QTgw.ipk-7D</i>	7D	
Vitesse de remplissage du grain	32 QTL	1A, 1B, 2A, 2D, 3A, 3B, 3D, 5B, 6D	Wang <i>et al</i> , 2009
Poids de 1000-grain	21 QTL	1A, 1B, 2A, 2D, 3B, 4A, 4D, 5A, 5B, 6D, 7D	
Poids total en grain sur l'épi	10 QTL	1A, 2A, 3B, 4B, 4D, 6B	
Longueur du grain	8 QTL	1A, 2D, 3A, 3B, 5A, 6A, 7A, 7D	
Poids de 1000-grain	10 QTL	1B, 2A, 2B, 3A, 3B, 4B, 4D, 5A, 6A, 7A	
Largeur du grain	9 QTL	1B, 1D, 2A, 3A, 3B, 4B, 5A, 6A, 7B	Gegas <i>et al</i> , 2010
Nombre de talles fertiles	1 QTL, <i>QTn.mst-6B</i>	6B	
			Naruoka <i>et al</i> , 2011

3.1 Exemples de QTL de rendement en grains identifiés chez le blé

Chez le blé, de nombreux QTLs contrôlant le rendement en grains et ses composantes ont été génétiquement identifiés (McCartney *et al*, 2005 ; Kuchel *et al*, 2007 ; Gegas *et al*, 2010 ; Tableau 3). Des QTLs pour le nombre de talles, le nombre de talles fertiles et le nombre d'épis ont été identifiés sur plusieurs chromosomes de blé (Richards, 1988 ; Shah *et al*, 1999 ; Kato *et al*, 2000 ; Li *et al*, 2002 ; Huang *et al*, 2003 ; An *et al*, 2006 ; Narashimhamoorthy *et al*, 2006 ; Kumar *et al*, 2007 ; Deng *et al*, 2011 ; Naruoka *et al*, 2011). Au cours de la dernière décennie, de nombreuses études ont reporté des QTLs pour le PMG et la taille du grain sur pratiquement l'ensemble des chromosomes du blé (Börner *et al*, 2002 ; Dholakia *et al*, 2003 ; Groos *et al*, 2003 ; Quarrie *et al*, 2005 ; Kumar *et al*, 2006 ; Breseghello et Sorrells, 2007 ; Li *et al*, 2007 ; Sun *et al*, 2009 ; Wang *et al*, 2009 ; Gegas *et al*, 2010 ; Ramya *et al*, 2010). Devant le très grand nombre d'études, seuls quelques exemples seront cités ici. Les allèles favorables aux loci *QYld.crc-2B* et *QYld.crc-4A*, situés respectivement sur les chromosomes 2B et 4A, augmentent par exemple de 0,18 tonnes.hectare⁻¹ et 0,12 tonnes.hectare⁻¹ respectivement la production de grains dans la population RL4452 x 'AC Domain', dans 6 environnements testés (McCartney *et al*, 2005). Le QTL *QTgw.ipk-7D*, sur le chromosome 7D, impacte de 10% le PMG dans des lignées d'introgression de blé synthétique W-7984 dans le fond génétique 'Prinz' (Röder *et al*, 2008). Un QTL sur le chromosome 1B, *QGyld.agt-1B*, a été détecté sur la population Trident x Molineux dans 18 environnements, et l'allèle favorable au locus augmente de 4,8% le rendement en grains (Kuchel *et al*, 2007).

Cependant, bien que le nombre de QTL identifiés pour des composantes du rendement soit très important, très peu de gènes contrôlant le rendement ont été clonés chez le blé tendre et aucun gène dans un QTL contrôlant le PMG ou la taille du grain de blé n'a été à ce jour cloné et caractérisé fonctionnellement (Yan *et al*, 2000, 2004 ; Fu *et al*, 2009).

De manière contrastée, chez le riz, plusieurs gènes contrôlant des composantes du rendement ont été isolés et étudiés fonctionnellement (Xing et Zhang, 2010).

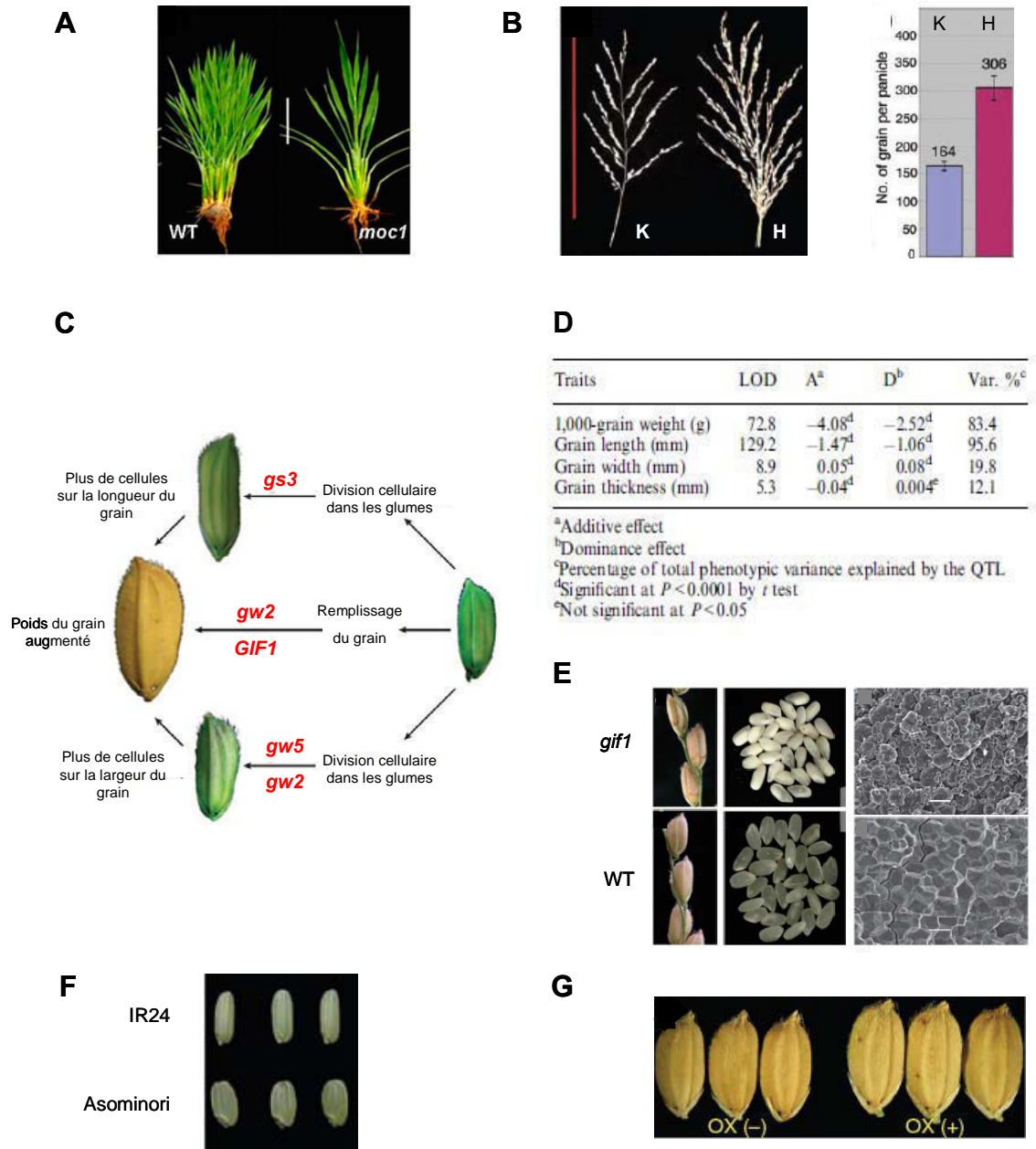


Figure 27 : Exemples de gènes contrôlant des composantes du rendement isolés et caractérisés fonctionnellement chez le riz. (A) Comparaison de l'habileté au tallage entre une variété mutante *moc1* et une variété sauvage (WT ; Li *et al*, 2003). (B) Comparaison de la structure de la panicule (panneau gauche) et du nombre de grains par panicule (panneau de droite) entre la variété Koshihikari (K) et la variété Habataki (H), qui porte l'allèle favorable (non-fonctionnel) de *Gn1a* (Ashikari *et al*, 2005). (C) Représentation schématique de l'effet des gènes *GS3*, *GIF1*, *GW5* et *GW2* dans le contrôle de la taille du grain chez le riz (Xing et Zhang, 2010). (D) Effet du QTL *GS3* sur la taille et le poids du grain de riz, détecté dans une population dérivée d'un backcross (BC3F1) de parents Minghui 63 et Chuan 7 (Fan *et al*, 2006). *GS3* est un QTL majeur pour le poids et la longueur du grain et un QTL mineur pour la largeur et l'épaisseur du grain. (E) Comparaison de grains avec (panneau gauche) et sans glumelles (panneau du milieu) entre une variété sauvage (WT) et une variété mutante *gif1* (Wang *et al*, 2008). Observation en microscopie électronique à balayage des granules d'amidon de grains sauvages et mutants *gif1* (panneau de droite). (F) Comparaison de grains entre la variété IR24 et la variété Asominori, portant l'allèle favorable (non-fonctionnel) de *GW5* (Weng *et al*, 2008). (G) Comparaison des grains entre des plants transgéniques sur-exprimant *GS5* (OX +) et des plants null-ségrégants (OX -) (Li *et al*, 2011).

3.2 Exemples de gènes de composantes du rendement en grains isolés et caractérisés chez le riz

Le gène *MOC1* (*Monococulm 1*) est un régulateur clé du nombre de talles chez le riz (Li *et al*, 2003 ; Figure 27, A). Ce gène a été isolé dans un crible de mutants de tallage et est situé sur le bras long du chromosome 6 du riz. Le gène *MOC1* code pour un facteur de transcription de la famille GRAS, spécifique des plantes et contrôle l'initiation et le développement des bourgeons axillaires, à l'origine des talles. Le mutant *moc1* présente un plateau de tallage sans talles, causé par un défaut de formation des méristèmes axillaires. La protéine MOC1 contrôlerait également la hauteur de la plante, puisque des plants de riz transgéniques surexprimant le gène *MOC1* présentent une taille réduite.

Le gène *Gn1a* (*Grain number 1 a*) a été isolé dans un QTL pour le nombre de grains, sur le chromosome 1 du riz (Ashikari *et al*, 2005 ; Figure 27, B). *Gn1a* code pour une cytokininine oxydase/déshydrogénase (OsCKX2), enzyme qui dégrade la phytohormone cytokininine. Une expression réduite d'OsCKX2 entraîne une accumulation de la cytokininine dans les méristèmes inflorescentiels, augmentant par conséquent le nombre d'organes reproducteurs et donc le rendement en grains, sans affecter la phénologie de la plante.

Les gènes *GS3*, *GIF1*, *GW5*, *GS5* et *GW2* ont été isolés dans le contrôle de la taille du grain de riz (Figure 27, C). *GS3* (*Grain Size 3*) a été isolé dans un QTL majeur pour la longueur et le poids du grain et mineur pour la largeur et l'épaisseur du grain, en région centromérique du chromosome 3 (Fan *et al*, 2006 ; Figure 27, D). L'allèle favorable de *GS3* permet une augmentation du poids du grain de riz de 50%, augmentation majoritairement attribuée à une augmentation de la longueur du grain causée par une augmentation du nombre de cellules sur l'axe longitudinal du grain. Le gène *GS3* code pour une protéine transmembranaire dont la fonction est encore inconnue. *GIF1* (*GRAIN INCOMPLETE FILLING 1*) a été isolé dans un crible de mutants de remplissage du grain de riz et est localisé sur le chromosome 4 du riz (Wang *et al*, 2008). Le gène *GIF1* code pour une invertase pariétale nécessaire à la redistribution du carbone au cours de la phase de remplissage du grain. *GIF1* est un régulateur positif du remplissage en réserves du grain puisque les grains du mutant perte-de-fonction *gif1* se remplissent moins vite, sont moins remplis et plus friables par rapport aux grains de variétés sauvages (Figure 27, E). Les granules d'amidon des grains de ces mutants sont anormalement développées et plus lâchement compactées, conduisant à une forte réduction du poids du grain. Le gène *GW5* (*Grain Weight 5*, également appelé *qSW5*) a été identifié dans un QTL majeur pour le poids et la largeur du grain sur le chromosome 5 du riz (Weng *et al*, 2008). *GW5* est un régulateur négatif de ces composantes

puisque des variétés possédant le gène *GW5* fonctionnel présentent des diminutions de 16,4% et 18,7% respectivement de la largeur et du PMG, par rapport aux mutants de délétion ne possédant pas *GW5* (Figure 27, F). Les grains des variétés possédant l'allèle mutant présentent des grains plus lourds du fait d'une augmentation du nombre de cellules dans les glumes qui autorise indirectement une augmentation de la largeur du grain. Le gène *GW5* code pour une protéine nucléaire de 144 acides aminés dont la fonction est pour l'instant inconnue. Cependant les auteurs ont montré que la protéine GW5 pouvait interagir physiquement avec des chaînes de polyubiquitine ; et suggèrent que GW5, tout comme GW2, interviendrait dans la régulation de la division cellulaire au cours du développement du grain de riz, via la voie de protéolyse ciblée Ubiquitine-Protéasome 26S. Récemment, le gène *GS5* (*Grain Size 5*) a été identifié dans un QTL majeur pour la largeur, le poids et le remplissage du grain sur le chromosome 5 du riz (Li *et al*, 2011). Le gène *GS5* code pour une sérine carboxypeptidase et est un régulateur positif de la taille du grain puisque une augmentation de l'expression de *GS5* est positivement corrélée à la largeur du grain (Figure 27, G).

Le gène *GW2* (*Grain Weight 2*) a été isolé dans un QTL majeur pour le poids et la largeur du grain (Song *et al*, 2007). Ce travail de thèse ayant pour objet l'analyse fonctionnelle de l'homologue de *GW2* chez le blé tendre, une section détaillée (section 7) est entièrement consacrée au gène *GW2* chez le riz et ses homologues en p.45 de cette synthèse bibliographique.

Comme nous venons de le voir avec les gènes *GW2* et *GW5*, l'ubiquitination et la voie de protéolyse ciblée lui étant associée, la voie UPS, constituent une des voies de signalisation régulant des composantes du rendement chez le riz. La voie UPS a largement été décrite, avec les facteurs de transcription et la phosphorylation des protéines, comme une des voies majeures de régulation des processus cellulaires, dans les cellules végétales. Pour ces raisons, au sein de l'équipe Rendement et Adaptation du Blé aux contraintes abiotiques, nous nous intéressons aux mécanismes moléculaires régissant le développement du grain en conditions optimales, et plus particulièrement à la signalisation par l'ubiquitination dans ce contexte scientifique.

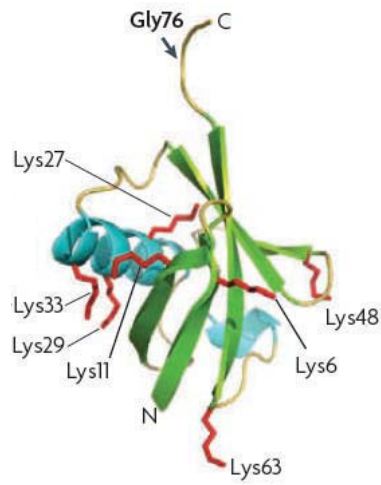


Figure 28 : Structure tridimensionnelle de l'ubiquitine (d'après Vierstra, 2009). N : extrémité N-terminale ; C : extrémité C-terminale. Les feuillets β sont représentés en vert et les hélices α en bleu. L'ubiquitine possède 7 résidus lysine (en rouge), et une glycine (Gly76) à l'extrémité C-terminale.

4 L'ubiquitination

L'ubiquitination est une modification post-traductionnelle réversible qui consiste en la fixation covalente d'une ou de plusieurs protéines d'ubiquitine sur une ou plusieurs lysines acceptrices d'une protéine substrat.

4.1 L'ubiquitine

L'ubiquitine est une protéine de 76 acides aminés (8,5 kDa), présente chez tous les eucaryotes. Elle est extrêmement conservée entre les règnes animaux et végétaux : sa séquence protéique primaire est strictement identique chez tous les végétaux supérieurs et diffère de deux et trois acides aminés chez la levure et les espèces animales, respectivement (Callis *et al*, 1995). La structure tridimensionnelle de l'ubiquitine est une forme globulaire très compacte, composée de 5 feuillets bêta formant une cavité au sein de laquelle vient se placer une hélice alpha (Figure 28). De nombreuses liaisons hydrogène intramoléculaires confèrent à l'ubiquitine une remarquable stabilité, sans doute essentielle pour prévenir sa dégradation au cours des cycles d'ubiquitination. L'ubiquitine possède une extrémité C-terminale dépassant de la structure globulaire de la protéine, terminée par une glycine (glycine 76).

L'ubiquitine est codée par la famille multigénique *UBQ*, composée de 16 gènes chez *Arabidopsis thaliana* (revue de Smalle et Vierstra, 2004). Les membres de cette famille *UBQ* expriment l'ubiquitine soit sous forme de polymères d'ubiquitine (gènes *polyUb*), dans lesquels plusieurs unités codantes d'ubiquitine de 228 pb sont organisées en concatémères, soit en fusion N-terminale d'une autre protéine (gènes *Ub-fusion*). Les gènes *polyUb* codent pour un nombre variable de répétition d'ubiquitine concaténées : chez *Arabidopsis* par exemple, il existe des gènes *polyUb* codant 3, 4, 5 et 6 unités d'ubiquitine à la suite ; chez le tournesol et le maïs, il existe 2 gènes codant un concaténaire de 6 répétitions d'ubiquitine et 2 gènes de 7 répétitions d'ubiquitine, respectivement. Les gènes *Ub-fusion* codent pour une ubiquitine en fusion N-terminale des protéines ribosomales S27a et L40 (Finley *et al*, 1989 ; Redman *et al*, 1989). Quelque soit le type de précurseur d'ubiquitine, les monomères d'ubiquitine sont libérés à partir des précurseurs traductionnels par les enzymes déubiquitinases, ou DUB, qui coupent spécifiquement la liaison peptidique qui suit la glycine C-terminale (glycine 76) de chaque monomère d'ubiquitine.

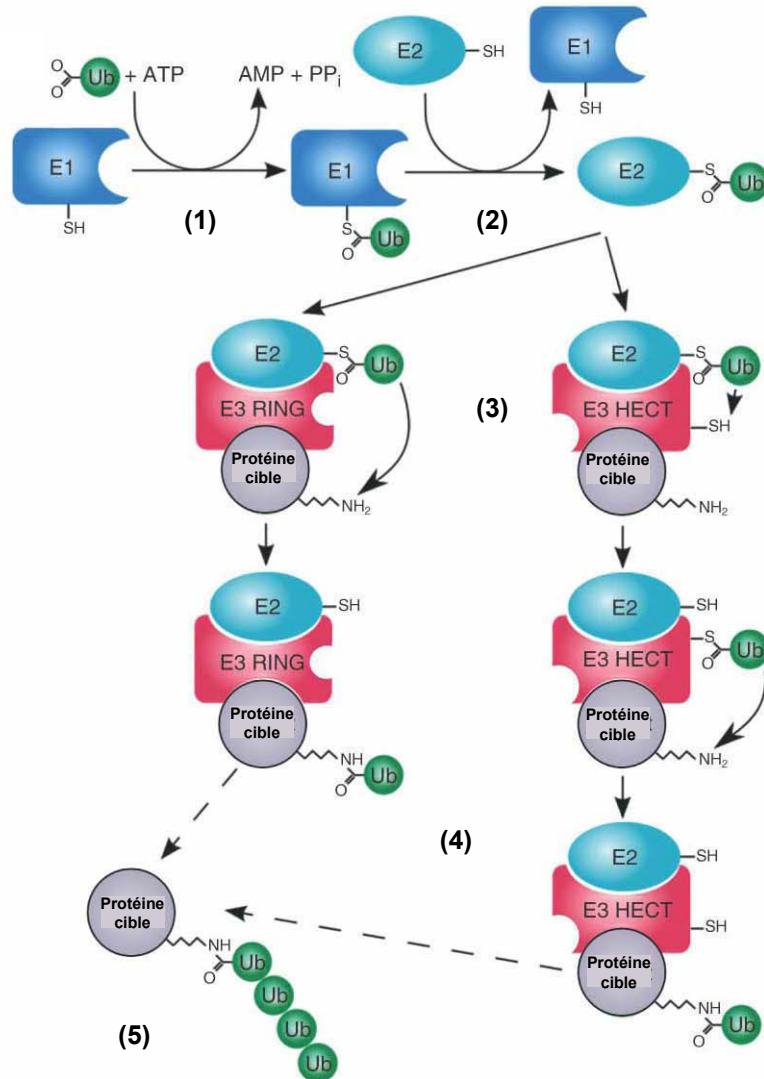


Figure 29 : Représentation schématisée simplifiée de la réaction d'ubiquitination (d'après Brooks, 2010). La réaction d'ubiquitination consiste en l'attachement covalent d'une ubiquitine (Ub) sur une protéine cible et nécessite l'intervention séquentielle des enzymes E1, E2 et E3. Suivant une réaction dépendante de l'ATP, l'enzyme E1 active l'ubiquitine (1), puis l'ubiquitine activée est transférée sur l'enzyme E2 (2). L'enzyme E3 reconnaît la protéine à étiqueter, et fixe l'enzyme E2 (3). L'ubiquitine est ensuite soit transférée directement sur la protéine cible (cas de la E3 RING), soit transférée sur l'enzyme E3 avant d'être transférée sur la protéine ciblée (cas de la E3 HECT) (4). Le produit final de cette réaction est la protéine substrat étiquetée avec un monomère d'ubiquitine. Plusieurs cycles réitérés de cette réaction conduisent à l'attachement d'une chaîne d'ubiquitine (polyubiquitination) sur la protéine cible (5).

Les promoteurs *UBQ* sont très actifs et ont notamment été exploités pour l'expression de transgènes chez les céréales.

4.2 La réaction enzymatique d'ubiquitination

La réaction d'ubiquitination fait intervenir successivement 3 enzymes : l'enzyme d'activation de l'ubiquitine, E1 ; l'enzyme de conjugaison E2 et l'enzyme E3 ligase (Figure 29). La réaction commence avec la formation dépendante de l'ATP d'une liaison thioester entre une cystéine de l'enzyme E1 et la glycine 76 de l'ubiquitine. L'ubiquitine ainsi activée est transférée sur une cystéine de l'enzyme E2 par transestérification. De son côté, l'enzyme E3 ligase reconnaît spécifiquement une protéine substrat. L'enzyme E2 se fixe ensuite à l'enzyme E3 et lie soit directement l'ubiquitine sur le groupement amine d'une lysine de la protéine substrat ou transfère l'ubiquitine sur l'enzyme E3 qui catalysera ensuite le transfert de l'ubiquitine sur une lysine de la protéine ciblée. Le produit final de cette réaction est un conjugué protéine substrat-ubiquitine dans lequel le groupement carboxyle de la glycine 76 de l'ubiquitine est liée par une liaison isopeptidique au groupement amine d'une lysine de la protéine substrat.

4.3 Les divers types d'ubiquitination

La protéine substrat peut être ubiquitinée de différentes manières ce qui conditionne son devenir dans la cellule. Une seule ubiquitine ou des chaînes de poly-ubiquitine peuvent être attachées sur un résidu du substrat (mono-ubiquitination et poly-ubiquitination, respectivement) ou des monomères d'ubiquitine sur plusieurs résidus du substrat (multi-mono-ubiquitination), au niveau d'une lysine le plus souvent (revue de Haglund et Dikic, 2005 ; Figure 30). Dans le cas de la poly-ubiquitination, l'élongation de la chaîne de poly-ubiquitine se fait par la formation d'une liaison isopeptidique entre la glycine 76 d'une nouvelle ubiquitine et une lysine de la molécule d'ubiquitine précédemment conjuguée au substrat, par des cycles successifs d'ubiquitination précédemment décrits.

L'ubiquitine contient 7 lysines (K6, K11, K27, K29, K33, K48, K63 ; Figure 28). Toutes les lysines sont exposées au solvant, et différents types de chaînes poly-ubiquitine peuvent ainsi être formées en fonction de la lysine de l'ubiquitine impliquée dans la liaison avec l'ubiquitine suivante dans la chaîne (revue par Trempe, 2011). Trois natures de chaînes

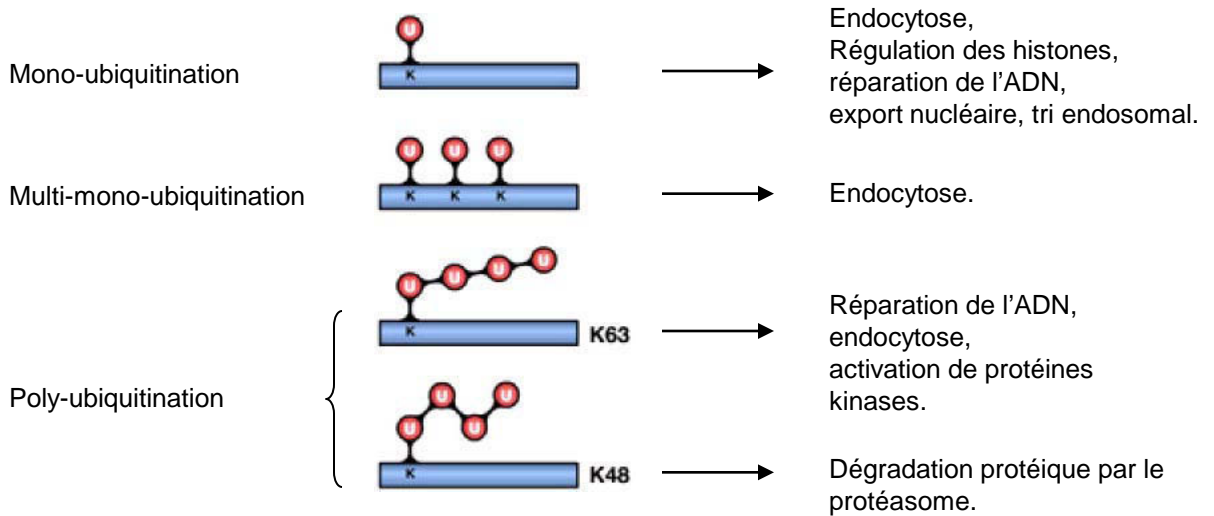


Figure 30 : Les différentes formes d'ubiquitination et leurs rôles (d'après Haglund et Dikic, 2005). Dans le cas de la poly-ubiquitination, seule le rôle des types de poly-ubiquitination le plus fréquemment rencontré (chaînes d'ubiquitine en K48 et K63) est mentionné.

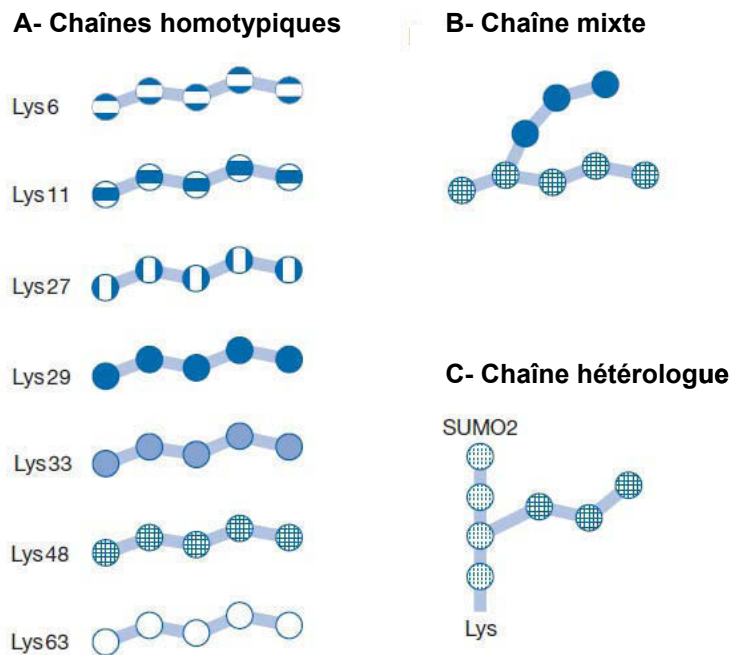


Figure 31 : Représentation schématique des différentes possibilités de chaînes d'ubiquitine (d'après Ikeda et Dikic, 2008). (A) Les chaînes homotypiques sont constituées d'ubiquitine reliées entre elles par une liaison incluant toujours la même lysine. (B) Les chaînes mixtes présentent des liaisons incluant diverse lysines des ubiquitines et sont bifurquées. (C) Les chaînes hétérologues sont l'assemblage d'ubiquitine et de protéine ubiquitin-like, comme la protéine SUMO 2 (Small Ubiquitin-like Modifier 2) par exemple.

peuvent être rencontrées : les chaînes homotypiques, majoritaires dans la cellule, dans lesquelles la même lysine est toujours impliquée dans la liaison entre 2 Ub au sein de la chaîne ; les chaînes hétérotypiques (ou chaînes mixtes) avec des liaisons impliquant différentes lysines et les chaînes hétérologues formées par l'assemblage d'Ub et de protéines Ubiquitin-like (Ikeda et Dikic, 2008 ; Trempe, 2011 ; Figure 31).

Le type de chaîne majoritairement formé est une chaîne d'ubiquitines reliées par la lysine 48 (chaînes K48) et conduit à la fonction la plus connue de l'ubiquitination : la dégradation de protéines par le protéasome 26S (Van Nocker et Vierstra, 1993 ; Trempe, 2011). L'apposition des autres types de chaînes polyUb et la monoubiquitination sont quant à elles impliquées dans des fonctions non-protéolytiques. Les chaînes en lysine 63 (chaînes K63) ont été montrées impliquées dans l'activation d'une protéine kinase humaine, dans la réparation de l'ADN et dans l'endocytose (Weissman, 2001 ; Haglund et Dikic, 2005). Un étiquetage par des chaînes en lysine 11 (chaînes K11) intervient dans le contrôle du cycle cellulaire et dans la voie de dégradation ERAD (Endoplasmic Reticulum-associated Degradation)/Proteasome. Les chaînes en lysine 6 (chaînes K6) ont été décrites dans la réparation de l'ADN (Wu-Baer *et al*, 2003 ; Sobhian *et al*, 2007). Bien qu'il ait été montré *in vivo* la formation des chaînes en lysine 27, 29 et 33 (K27, K29 et K33 respectivement), leur fonction est encore inconnue (Trempe, 2011). La monoubiquitination conduit à l'endocytose, la réparation de l'ADN, la régulation des histones, le tri endosomal, l'export nucléaire et le bourgeonnement viral Haglund et Dikic, 2005).

Ces différentes étiquettes peuvent ensuite être reconnues par des domaines protéiques appelés UBD (Ubiquitin-Binding Domain). A l'heure actuelle, 9 UBDs différents ont été identifiés (Haglund et Dikic, 2005).

4.4 Les enzymes déubiquitinases (DUBs)

Les DUBs interviennent à plusieurs étapes de la voie d'ubiquitination : elles génèrent d'une part les monomères d'ubiquitine à partir de leur précurseur comme mentionné précédemment, mais recyclent également les monomères d'ubiquitine après dégradation des protéines polyubiquitinées et sont capables d'inverser la réaction d'ubiquitination. Le génome d'*Arabidopsis* compte 64 gènes de déubiquitinases (Vierstra, 2009).

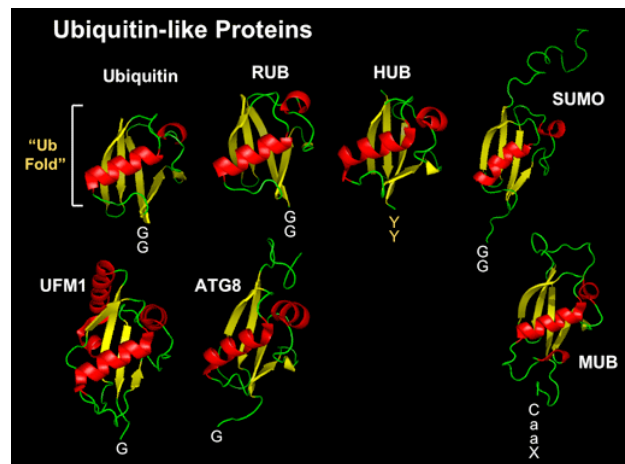


Figure 32 : Structure tridimensionnelle de protéines Ubiquitin-like (Site internet du Laboratoire de RD Vierstra). RUB, Related Ubiquitin ; HUB, Homologous to Ubiquitin ; SUMO, Small Ubiquitin-like Modifier ; UFM1, Ubiquitin-fold modifier 1 ; ATG8, Autophagy-defective 8 ; MUB, Membrane-anchored Ubiquitin fold.

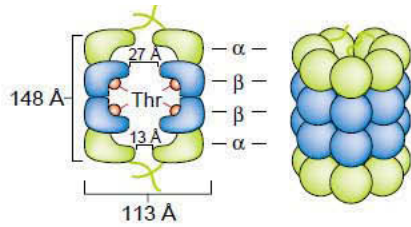
4.5 Les protéines Ubiquitin-like (UBLs)

Depuis la découverte de la protéine ubiquitine, d'autres protéines apparentées ont été identifiées (Smalle et Vierstra, 2004). Chez les plantes, ce sont entre autres les protéines RUB1 (Related Ubiquitin 1 ou NEDD8), SUMO (Small Ubiquitin-related Modifier), APG8 (Autophagy-defective 8), APG12 (Autophagy-defective 12), URM (Ubiquitin-Related Modifier) et HUB (Homologous to Ubiquitin). Bien qu'elles partagent peu d'homologie de séquence avec l'ubiquitine, toutes présentent la conformation tridimensionnelle globulaire de l'Ub avec une extension C-terminale flexible similaire (Figure 32). Tout comme l'ubiquitine, elles constituent une étiquette qui sera apposée sur une protéine cible selon une réaction dépendante de l'ATP impliquant des enzymes E1 et E2 et quelques fois une enzyme E3. Des protéines DUB-like peuvent également les retirer du substrat. A l'exception de SUMO qui peut être retrouvée attachée à la protéine cible sous forme de polymères, les UBLs sont attachées sous forme de monomère. Tout comme l'Ub, les conséquences de l'apposition d'UBL peuvent être diverses (Smalle et Vierstra, 2004). Par exemple, un étiquetage par RUB1 modifie de manière réversible l'activité de E3s de type SCF en se liant à la sous-unité Cullin (Smalle et Vierstra, 2004). Une modification de type SUMO peut affecter l'activité d'une protéine ou sa localisation subcellulaire, ou encore protéger des protéines d'une ubiquitination. Dans au moins 2 protéines de mammifère, une même lysine peut être soit sumoylée ou ubiquitinée ; la sumoylation bloquant la dégradation de la protéine par la voie Ubiquitine-Protéasome 26S. Kurepa et ses collaborateurs (2003) suggèrent qu'un tel mécanisme de protection de protéines contre la protéolyse par sumoylation existerait aussi chez les plantes, notamment en cas de stress.

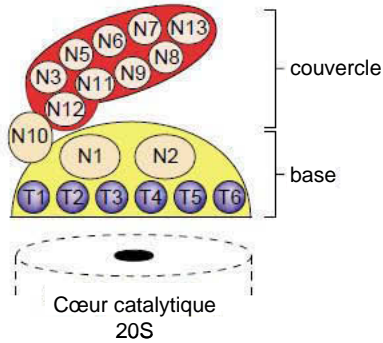
5 La voie de protéolyse ciblée Ubiquitine-Protéasome 26S.

La capacité des cellules à basculer d'un stade de développement à un autre ou encore à s'adapter à de nouvelles conditions environnementales demande de pouvoir rapidement synthétiser de nouveaux polypeptides et dégrader des régulateurs préexistants. Différents systèmes de dégradation plus ou moins spécifiques des protéines coexistent au sein de la cellule. Dans la mitochondrie, la vacuole et certains plastes, des protéases dégradent les protéines de manière non-spécifique, par reconnaissance de certains acides aminés ou de certains sites (comme par exemple les protéases à cystéine, à acide aspartique, à sérine, les

A- La partie catalytique, le protéasome 20S



B- Le complexe régulateur 19S



C- Le protéasome 26S

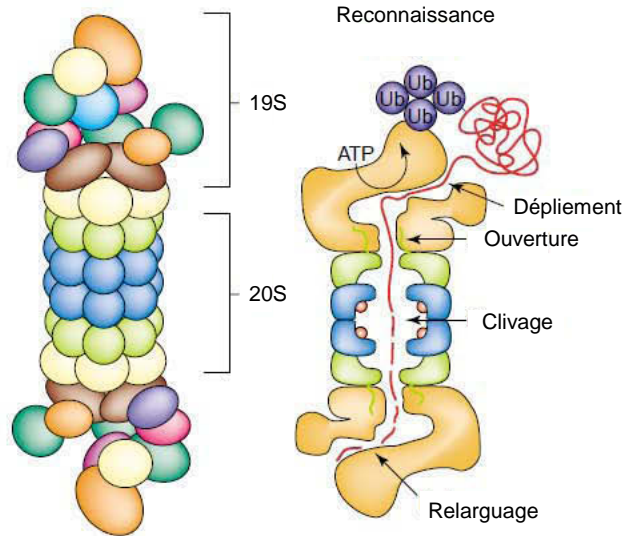


Figure 33 : Organisation et structure du protéasome 26S (d'après Vierstra, 2003). Organisation du protéasome 20S (A) et du complexe régulateur 19S (B). N, sous-unités non-ATPasiques ; T, sous-unités ATPases de type AAA. (C) Structure du protéasome 26S et séquence des événements conduisant à la dégradation d'une protéine polyubiquitinée en K48 par le protéasome 26S.

métalloprotéases, les endo- et exo-peptidases). La voie Ubiquitine-Protéasome 26S, ou UPS, est le seul système à cibler spécifiquement des protéines pour une dégradation. Il consiste en l'ubiquitination de protéines substrat par l'intervention successive des enzymes E1, E2 et E3, puis d'une reconnaissance et d'une dégradation des protéines polyubiquitinées par le protéasome 26S. Bien que ce système existe chez la levure et tous les animaux, il a majoritairement été privilégié chez les plantes où il est d'une importance et d'une complexité considérable. Chez *Arabidopsis thaliana*, les composants de cette voie représentent approximativement 5% du protéome (Smalle et Vierstra, 2004). Il est par conséquent considéré maintenant comme l'un des systèmes de régulation majeur des fonctions cellulaires chez les plantes, avec la transcription et la phosphorylation des protéines. La haute spécificité de ce système de dégradation est apportée par le nombre important d'enzymes E3 ligases. Le génome d'*Arabidopsis thaliana* compte par exemple plus de 1500 gènes de E3, contre seulement 2 gènes de E1 et 37 gènes de E2 (Vierstra, 2003 ; Hua et Vierstra, 2011). La reconnaissance E3 ligase-substrat a longtemps été imaginée comme un système clé-serrure ; le nombre de protéines ciblées étant tout aussi important qu'il existe de E3. Cependant, plusieurs exemples d'une même E3 pouvant cibler plusieurs substrats différents ont été récemment reportés, suggérant que le nombre de protéines régulées par la voie UPS serait bien plus important encore (Yi et Deng, 2005 ; Boulton, 2006).

5.1 Le protéasome 26S.

Le protéasome 26S reconnaît puis dégrade les protéines étiquetées avec une chaîne de polyubiquitine comportant au minimum 4 monomères d'ubiquitine, reliés en K48. Le protéasome 26S est un complexe protéolytique dépendant de l'ATP, de 2,5 MDa, présent dans le noyau et le cytoplasme des cellules animales et végétales. Il est fortement présent dans les tissus en division et lors de stress. Le protéasome 26S est constitué d'un cœur catalytique en forme de baril, le protéasome 20S ou CP (*core particle*), et de 2 complexes régulateurs 19S, ou RP (*regulatory particle*), placés à chacune des extrémités du cœur catalytique (Vierstra, 2003 ; Figure 33).

5.1.1 Le protéasome 20S.

Le 20S, ou CP (*core particle*), est un complexe de 720 kDa composé de 28 sous-unités arrangées en 4 anneaux heptamériques empilés : 2 anneaux de sous-unités α aux extrémités de

la structure et 2 anneaux de sous-unités β au centre, enfermant une chambre protéolytique interne (Figure 33, A). Le complexe 20S est une protéase à large spectre, indépendante de l'ATP. Les sites protéolytiques actifs au sein de la chambre interne du CP sont des thréonines N-terminales des sous-unités $\beta 1$, $\beta 2$ et $\beta 5$. Les sous-unités $\beta 1$, $\beta 2$ et $\beta 5$ ont respectivement des activités peptidylglutamyle hydrolase, trypsin-like et chymotrypsine-like, donnant ainsi au CP la capacité de cliver la plupart des liaisons peptidiques (Kurepa et Smalle, 2008). Le CP dégrade les polypeptides préalablement reconnus et dépliés par le RP, en peptides de 6 à 9 résidus.

Tout comme chez la levure et les animaux, les sites actifs du 20S chez les plantes sont sensibles aux inhibiteurs sélectifs chimiques MG115, MG132, lactacystine et epoxomycine (Yang *et al*, 2004).

5.1.2 Le complexe régulateur 19S.

Chaque 19S, ou RP (*regulatory particle*), est un complexe multimérique de 890 kDa. Chez *Arabidopsis*, le 19S est composé de 17 sous-unités pouvant être divisées en deux sous-complexes : le couvercle et la base (Glickman et Raveh, 2005 ; Kurepa et Smalle, 2008 ; Figure 33, B). La base est en contact direct avec le 20S. Elle est composée d'un anneau de 6 ATPases de type AAA (sous-unités RPT 1 à 6) et de 3 sous-unités non-ATPasiques (RPN1, 2 et 10). Les sous-unités RPT utilisent l'hydrolyse de l'ATP pour aider au dépliement des polypeptides ciblés et ouvrir l'accès à la chambre interne du 20S (Hartmann-Petersen *et al*, 2003). Le couvercle du 19S est constitué de 8 sous-unités non ATPasiques (RPN3, 5 à 9 et RPN 11 et 12). Les fonctions du 19S sont la reconnaissance et la fixation de la protéine polyubiquitinée en K48, le retrait de la chaîne de polyUbs, le dépliement de la protéine cible et le transfert dans la chambre protéolytique du 20S. Les peptides clivés sortent du 20S par l'ouverture du cylindre grâce à l'énergie produite par l'hydrolyse de l'ATP catalysée par les sous-unités ATPasiques de la base du 19S.

D'autres protéines peuvent se lier plus faiblement au complexe 26S, suggérant que ce complexe multimérique serait à l'origine d'un plus grand complexe dynamique (Schmidt *et al*, 2005). Par exemple, chez les plantes, 2 protéines kinases (un homologue de SNF1 et une protéine de la famille des kinases dépendante du calcium) interagissent avec le 26S sans doute pour réguler son activité (Farras *et al*, 2001 ; Lee *et al*, 2003). Chez la levure, la E3 de type HECT Huf5, la déubiquitinase Ubp6 et la protéine Ecm29 aident à l'interaction du RP avec le CP. Des homologues de ces gènes ont notamment été retrouvés chez *Arabidopsis*.

Tableau 4 : Exemples de la fonction de gènes de la voie UPS, dans le développement d'*Arabidopsis thaliana* (d'après Vierstra, 2003).

Gene	Protein type	function
UFO	E3 (F-box)	Flower development
TIR1	E3 (F-box)	Auxin responses
COI1	E3 (F-box)	Jasmonate responses
FKF1/ZTL/LKP2	E3 (F-box)	Circadian rhythms
EID1	E3 (F-box)	Photomorphogenesis
ORE9/MAX2	E3 (F-box)	Senescence
SON1	E3 (F-box)	Pathogen resistance
SKP2;1	E3 (F-box)	E2F degradation
ASK1	E3 (SKP)	Auxin responses
CUL1	E3 (Cullin)	Embryogenesis
RBX1	E3 (RBX)	Auxin responses
SGT1	E3 (SGT1–SCF)	Pathogen resistance
COP10	E2-like	Photomorphogenesis
COP1	E3 (Ring HC)	Photomorphogenesis
CER3	E3 (Ring HC)	Wax biosynthesis
PRT1	E3 (Ring HC)	N-end rule substrates
SINAT5	E3 (Ring HC)	Auxin responses
UPL3	E3 (HECT)	Trichome development
HOBBIT	E3 (APC)	Cell division
RPN12a	26S proteasome lid	Cytokinin responses
RPN10	26S proteasome base	ABA responses
UBP1 and 2	DUB	Abnormal protein degradation
UBP14	DUB	Embryogenesis

De plus, le protéasome 20S peut également être retrouvé seul dans les cellules. Il assure une dégradation indépendante de l'ubiquitine des protéines oxydées par les espèces activées de l'oxygène et jouerait un rôle important dans les mécanismes de défense contre les potyvirus TMV et LMV du fait de son activité RNAsique (Kurepa et Smalle, 2008).

5.2 Les fonctions de l'UPS

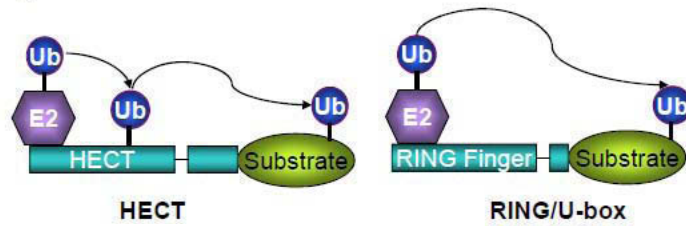
La voie UPS a lieu dans le noyau et le cytoplasme, mais également sur la surface interne de la membrane plasmique et dans le réticulum endoplasmique. Elle est utilisée pour dégrader des régulateurs clés à courte durée de vie mais sert également à dégrader des protéines anormales ou à mauvaise conformation. On estime que pas moins de 30% des produits initiaux de traduction sont non-fonctionnels et sont rapidement dégradés par la voie UPS (Schubert *et al*, 2000). Les régulateurs clés incluent notamment des enzymes limitantes de certaines voies métaboliques. En conférant une courte demi-vie à ces protéines, la voie UPS régule finement des flux métaboliques et peut diminuer le métabolisme quand le substrat est limité ou que le produit de la réaction métabolique est en excès et n'est plus nécessaire. Chez le soja, par exemple, la voie UPS dégrade une pyruvate kinase afin de contrôler la répartition du carbone dans la plante (Tang *et al*, 2003). Parmi les régulateurs clés sont également retrouvés des récepteurs de signalisation ainsi que des activateurs ou répresseurs de la transcription (Clough *et al*, 1999 ; Hellmann et Estelle, 2002 ; Vierstra, 2003).

La voie Ubiquitine-Protéasome 26S impacte presque tous les aspects de la croissance et du développement de la plante, affectant notamment le cycle cellulaire, l'embryogénèse, la photomorphogénèse, la sénescence, les mécanismes de défense, les réponses à des stimuli de l'environnement et la signalisation hormonale (Vierstra, 2009 ; Tableau 4).

6 Les enzymes E3

A ce jour, il existe 4 types principaux d'E3 ligases chez les plantes, classés en fonction de leur mécanisme d'action et de leur composition : les E3 HECT, RING, U-box et Cullin-RING ligases (CRLs). Seule cette dernière classe est représentée par des E3 multimériques et

A- Les E3 monomériques



B- Les E3 multimériques

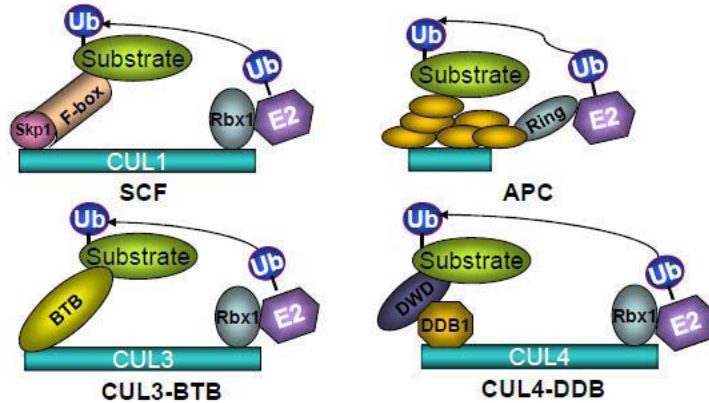


Figure 34 : Les différents types de E3 ligases (d'après Liu *et al*, 2012). Les HECT, les RING et les U-box sont des E3 monomériques ; tandis que les CRLs sont des E3 multimériques. La famille des E3 de type CRL comprend les complexes SCF, CUL4-DDB, CUL3-BTB et APC. Seules les HECT forment un intermédiaire E3-Ub avant de transférer l'ubiquitine (Ub) sur un résidu lysine de la protéine cible (en vert).

peut être subdivisée en 4 sous-types : les SCF, les CUL4-DDB, les CUL3-BTB et les APC (Viestra, 2009 ; Figure 34).

6.1 Les E3 HECT

Les HECT (Homology to E6AP C Terminus) sont des E3 ligases monomériques. Elles sont facilement identifiables par la présence d'un domaine conservé de 350 acides aminés en partie C-terminale, retrouvé pour la 1^{ère} fois dans la protéine humaine E6-Associated Protein, une E3 ligase qui dégrade p53 (Huibregtse *et al*, 1995 ; Downes *et al*, 2003). Le génome d'*Arabidopsis* compte 7 gènes de E3 ligases de type HECT, appelées UPL1 à 7 (Moon et Estelle, 2004). Ce sont des protéines volumineuses : chez *Arabidopsis*, les 7 E3 HECT ont une taille qui varie de 96 à 405 kDa (Smalle et Vierstra, 2004). Les E3 HECT sont uniques dans la famille des E3 ligases car elles semblent être les seules à former un intermédiaire thiol-ester E3-Ub avant de transférer l'Ub sur le substrat. Le domaine HECT permet l'interaction avec la E2 et contient le site actif de type cystéine, accepteur de l'ubiquitine. Les E3 HECT possèdent souvent en amont du domaine HECT d'autres domaines d'interaction protéine-protéine : par exemple des domaines Armadillo, C2, Ub-associated (UBA), Ub-interacting motif (UIM), Ub-like (UBL) ou encore Ankyrin (Downes *et al*, 2003). Ces domaines joueraient un rôle dans la reconnaissance du substrat, dans la localisation subcellulaire et la régulation de la protéine HECT elle-même ou encore dans la liaison de l'ubiquitine (Downes *et al*, 2003). Bien que les fonctions des E3 HECT chez les mammifères soient connues (rôles notamment dans le trafic protéique intracellulaire, la réponse immunitaire, la prolifération (Metzger *et al*, 2012), leur fonction est beaucoup moins bien établie chez les plantes. La plus étudiée d'entre elles, UPL3, intervient dans le développement des trichomes en contrôlant l'endoréduplication (Downes *et al*, 2003). Une autre E3 HECT, UPL5, régule la sénescence foliaire par dégradation du facteur de transcription WRKY53 (Miao et Zentgraf, 2010).

6.2 Les E3 RING

Les E3 de type RING sont caractérisées par un domaine RING, identifié pour la première fois dans la protéine humaine RING1 (Really Interesting New Gene 1). Elles sont présentes chez les animaux, les levures, les plantes et les virus. Chez l'Homme, elles sont notamment responsables de nombreuses maladies en cas de mutations de ces gènes. Chez les plantes, les E3 de type RING représentent, avec les E3 de type SCF, la plus grande famille de

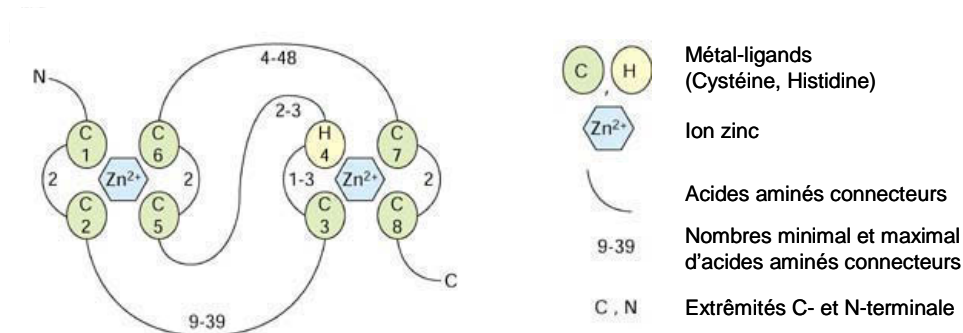


Figure 35 : Représentation schématique de la structure d'un domaine RING (d'après Kosarev *et al*, 2002). Le domaine RING est arrangement de 8 métal-ligands (résidus cystéine et histidine), chélatant 2 ions zinc. Les paires de métal-ligands 1 et 3 (métal-ligands 1-2 et 5-6, respectivement) chélatent un 1^{er} ion zinc tandis que les paires 2 et 4 (métal-ligands 3-4 et 7-8, respectivement) chélatent un 2^{ème} ion zinc.

E3. Le génome d'*Arabidopsis* compte 499 gènes de E3 RING (Mazzucotelli *et al*, 2006). Elles sont impliquées dans de nombreux processus physiologiques comme par exemple la photomorphogénèse, la signalisation hormonale (par l'acide abscissique et l'auxine), la perception du froid ou encore l'auto-incompatibilité.

Dans la très grande majorité des cas, les E3 à domaine RING sont monomériques (E3 RING). Elles possèdent à la fois le domaine RING, nécessaire à l'interaction avec l'enzyme E2, et le ou les domaine(s) nécessaire(s) à la reconnaissance de la protéine substrat. Quelques protéines RING peuvent également constituer une sous-unité des complexes Cullin-RING ubiquitine ligases (CRLs, cf. section 6.4). On retrouve également des E3 RING sous forme d'homodimère ou d'hétérodimère avec une autre protéine RING (Smalle et Vierstra, 2004). Contrairement aux E3 HECT, la plupart des E3 RING ne forme pas d'intermédiaire thioester avec l'ubiquitine. Elles agissent plutôt comme un échafaudage liant et rapprochant dans l'espace la E2 et le substrat protéique, et activent allostériquement l'enzyme E2. L'ubiquitine est alors transférée directement de la E2 à la protéine cible. Cependant, il semblerait qu'en absence de substrat protéique à étiqueter la E3 RING soit capable *in vitro* de transférer l'ubiquitine sur elle-même (Stone *et al*, 2005).

6.2.1 Le domaine RING

Le domaine RING est un domaine d'interaction protéine-protéine de 40 à 60 acides aminés. Ce domaine est un octet de résidus cystéine et histidine (appelés métal-ligands), arrangés en une structure arc-boutée croisée qui chélate 2 ions zinc (Figure 35). Le domaine RING est très similaire au motif d'interaction avec l'ADN en doigt de zinc ; cependant les motifs doigts de zinc consistent en 2 paires de métal-ligands chélatant 1 ion zinc contre 4 paires de métal-ligands chélatant 2 ions zinc dans le domaine RING (Kosarev *et al*, 2002).

Une analyse fonctionnelle des protéines à domaine RING chez *Arabidopsis* a permis d'identifier, selon la nature et la position des métal-ligands, 2 types de domaines RING canoniques, RING-H2 (C3H2C3, à savoir des cystéines en position 1 à 3 et 6 à 8, histidines en 4 et 5) et RING-HC (C3HC4), et 5 types de domaines RING modifiés (RING-C2, RING-v, RING-D, RING-S/T et RING-G) (Stone *et al*, 2005 ; Figure 36). Cependant cette classification pourrait varier entre les espèces végétales. En effet, le domaine RING modifié RING-D semble être caractéristique des dicotylédones puisqu'une recherche de similarité dans le génome du riz n'a pas permis à ce jour de retrouver des protéines avec un tel domaine.

RING Domains	No. of Domains	Consensus
		<div> <div>Zn2+</div> <div> <div>1</div> <div>2</div> <div>3</div> <div>4</div> <div>5</div> <div>6</div> <div>7</div> <div>8</div> </div> <div>metal ligand</div> </div>
RING-H2	241	C - x ₂ - C - x ₅₋₃₆ - C - x ₁₋₃ - H - x ₂₋₃ - H - x ₂ - C - x ₄₋₄₈ - C - x ₂ - C
RING-HC		C - x ₂ - C - x ₅₋₃₆ - C - x ₁₋₃ - H - x ₂₋₃ - C - x ₂ - C - x ₄₋₄₈ - C - x ₂ - C
RING-HCa	145	C - x - [I/V] - C - x _{10,16,20,28} - C - x ₁₋₂ - [G] - H - x ₁₋₂ - [F] - C - x ₂ - C - x ₃ - [W] - x ₆₋₃₀ - C - [P] - x - C
RING-HCb	41	C - x - [I/V] - C - x _{13,15,24} - C - x ₁₋₂ - [G] - H - x ₁₋₂ - [F] - C - x ₂ - C - x _{7,11,22} - C - [P] - x _{3,4,8} - C
RING-C2	10	C - x ₂ - C - x ₁₃ - C - x ₁ - C - x ₄ - C - x ₂ - C - x ₁₁₋₁₆ - C - x ₂ - C
		C - x - [I/L] - C - x ₃ - [DL] - x _{1,6} - C - x _{1,2} - C - x _{4,5} - C - x ₂ - C - x ₃ - [W/I] - x _{6,12} - C - [P] - x - C
RING-v	25	C - x ₂ - C - x ₁₀₋₄₅ - C - x ₁ - C - x ₇ - H - x ₂ - C - x ₁₁₋₂₄ - C - x ₂ - C
		C - [R I/V] - C - x _{10-45/67/69} - C - [S] - C - x ₆ - [A/V] - H - x ₂ - C - x ₃ - [W] - x ₃ - [K G] - x ₃₋₄ - C - [E I/V] - C
RING-D	10	C - x - [I/V] - C - [LEEF] - x ₂ - [G] - x ₅ - [LP] - C - [G] - H - x - [F] - D - x ₂ - C - x ₃ - [WF] - x ₄ - [H] - x - C - [PL] - C
RING-S/T	4	C - x - [V] - C/S - x _{14,15} - C - [G] - H - x - [F] - H/C - x _{2,3} - S/T - x _{2,3} - [W] - x ₇₋₁₁ - C - x ₂ - C
RING-G	1	C - x ₂ - C - x ₈₈ - C - x ₁ - H - x ₇ - G - x ₂ - C - x ₁₃ - C - x ₂ - C
		C - x - [I] - C - x ₅ - C - x ₁ - H - x - [Y] - G - x ₂ - C - x ₃ - [W] - x ₃ - C - [P] - x - C

Figure 36 : Séquences consensus et nombre de chaque type de domaines RING identifiés chez Arabidopsis (d'après Stone *et al*, 2005). Les domaines RING-H2 (C3H2C3) et RING-HC (C3HC4) sont des domaines canoniques ; les domaines RING-C2 (C8), RING-v (C4HC3), RING-D (C3HDC3), RING-S/T et RING-G (C3HGC3) sont des domaines RING modifiés.

Tous les types de domaines RING précédemment cités, excepté le domaine RING-S/T, ont été montrés comme étant fonctionnels en conditions *in vitro* (Stone *et al*, 2005).

Le domaine RING des E3 ligases permet l'interaction avec l'enzyme E2 (Deshaies et Joazeiro, 2009) et son intégrité est dans la grande majorité des cas essentielle pour l'activité E3 ligase. La nature de cette interaction RING-E2 a été étudiée pour des E3 RING humaines, par résonance magnétique nucléaire. Bien qu'il soit bien établi que ce domaine recrute l'enzyme E2, il existe une faible corrélation entre l'activité ubiquitine ligase et la capacité d'une E3 RING à lier une E2 avec une forte affinité. Par exemple, l'hétérodimère BRAC1/BARD1 peut stablement lier la E2 UbcH7 mais est inactif pour le transfert de l'ubiquitine (Christensen *et al*, 2007). D'autre part, des paires E2-E3 hautement actives ne présentent pas d'association stable. L'affinité de domaines RING isolés pour une E2 est généralement forte, si l'on prend, en première approximation, un K_m de l'ordre de quelques μM .

En plus des 8 acides aminés métal-ligands, d'autres régions ou d'autres résidus ont été montrés comme étant importants pour que le domaine RING permette l'activité E3 ligase. Par exemple, la région BRR (basic rich region) en amont du domaine RING de Ubr1 est essentielle pour la liaison à l'E2 mais seul le domaine RING permet l'activité ubiquitine ligase (Deshaies et Joazeiro, 2009). Dans plusieurs E3 RING comme Cnot4, c-Cbl et BRCA1, l'acide aminé précédant le 2^{ème} métal-ligand doit être un acide aminé hydrophobe, comme une leucine, une isoleucine ou une valine. Cet acide aminé est requis pour la liaison de la E2 et l'activité E3 puisque sa mutation abolit totalement l'activité (Zheng *et al*, 2000 ; Albert *et al*, 2002 ; Brzovic *et al*, 2003 ; Stone *et al*, 2005). De manière intéressante, cet acide aminé est conservé dans les 7 types de domaines RING décrits chez Arabidopsis.

6.2.2 Les domaines protéiques associés au domaine RING.

Les protéines à domaine RING possèdent souvent, en plus du domaine RING, d'autres motifs protéiques connus. Ceux-ci peuvent agir comme régulateur de la fonction de la protéine ou être des domaines d'interaction protéine-protéine nécessaires à la reconnaissance de la protéine cible, tels que des motifs coiled-coil, doigt de zinc (de type Cx2Cx5Cx2C), WD40 ou encore des répétitions ankyrins (Mazzucotelli *et al*, 2006). Dans quelques cas, le domaine RING est retrouvé associé également à des motifs de liaison à l'ADN tels que les motifs C2H2, C3H1, K Homology (KH) ou encore RNA Recognition Motif (RRM) (Mazzucotelli *et al*, 2006).

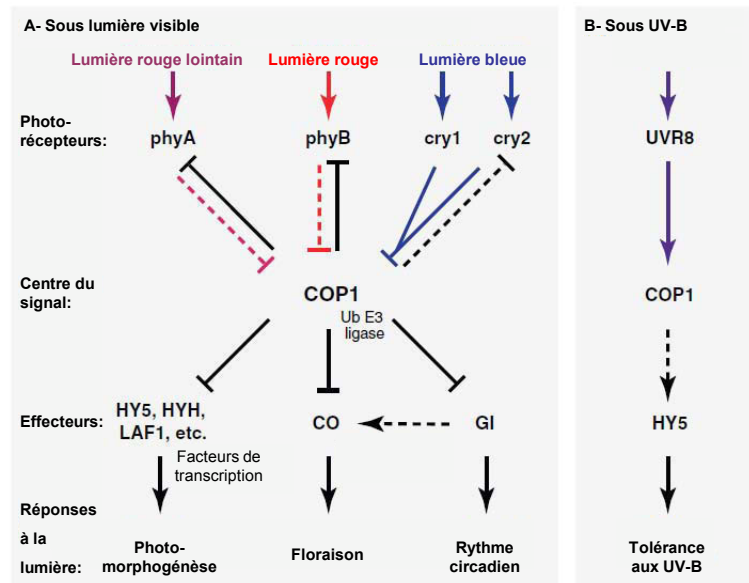


Figure 37 : COP1 est un régulateur central de la transduction du signal de réponse de croissance et de développement à la lumière (d'après Lau et Deng, 2012). COP1 centralise et intègre les signaux venant des différents photorécepteurs et contrôle en aval les diverses voies de réponse à la lumière. **(A)** COP1 est une E3 RING ligase qui ubiquitine des effecteurs positifs de la réponse à la lumière pour dégradation. Les phytochromes (phyA et phyB) et les cryptochromes (cry1 et cry2) sont les photorécepteurs majeurs, percevant un large spectre de la lumière visible. Lorsqu'ils sont activés, les photorécepteurs répriment la fonction de COP1, permettant l'accumulation des effecteurs et la mise en place de la réponse de développement à la lumière. **(B)** En présence d'UV-B, cependant, COP1 agit comme un régulateur positif de la voie de signalisation. Une fois activé, le récepteur des UV-B, UVR8 interagit avec COP1, ce qui permet l'activation de la transcription de HY5 et engendre la tolérance aux UV-B. Les traits pleins représentent une interaction physique démontrée pour la régulation du processus. CO, CONSTANTS; GI, GIGANTEA.

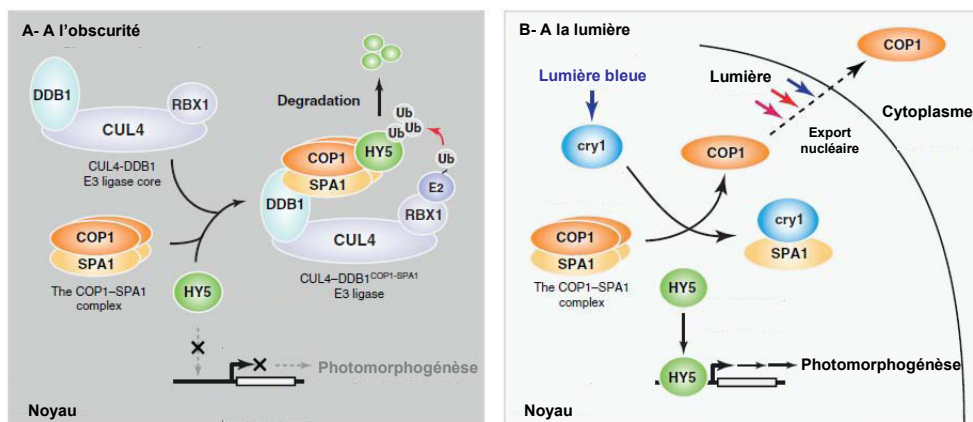


Figure 38 : Mode d'action de COP1 dans la dégradation de HY5 en réponse à la lumière visible (d'après Lau et Deng, 2012). **(A)** A l'obscurité, COP1 forme un complexe tétramérique avec SPA1, puis un super-complexe avec CUL4-DDB1 pour ubiquitiner et dégrader HY5, réprimant ainsi la photomorphogénèse. **(B)** A la lumière, COP1 est inhibée par les photorécepteurs activés. Sous lumière bleue, cry1 interagit et séquestre SPA1, inactivant COP1. HY5 s'accumule et permet la photomorphogénèse. A long terme, la lumière permet l'export nucléaire de COP1.

La combinaison de ces domaines avec un domaine RING peut être dans quelques cas plante-spécifique. Par exemple, les domaines vWA (von Willebrand factor (vWF) type A), KISc (Kinesin motor domain) et Znf-ZZ (Zinc finger) sont retrouvés dans diverses espèces mais leur association avec un domaine RING n'est retrouvée que chez les plantes (Stone *et al*, 2005).

6.2.3 Exemples d'E3 RING monomériques décrites chez les plantes.

6.2.3.1 La protéine COP1

La protéine COP1 (COnstitutive Photomorphogenic 1) est la E3 de type RING ayant été le plus étudiée chez les plantes. Elle a tout d'abord été identifiée dans un crible de mutants pour la photomorphogénèse. Les mutants *cop1* placés à l'obscurité présentaient un phénotype de plantules ayant poussé à la lumière (phénotype constitutif dé-étiolé), c'est-à-dire un hypocotyle court, un développement des feuilles et une activité photosynthétique (Deng *et al*, 1991).

COP1 est un intégrateur et un répresseur central de la voie de signalisation de croissance à la lumière visible chez les plantes (Figure 37). Dans cette voie, COP1 ubiquitine et amène à la dégradation des régulateurs positifs de cette voie et est elle-même régulée par de multiples photorécepteurs (Lau et Deng, 2012). Outre son rôle dans la photomorphogénèse, COP1 intervient également dans de nombreux autres processus, dont la plupart sont des réponses de croissance et de développement régulées par la lumière : la floraison, le rythme circadien, la signalisation aux UV-B, l'ouverture stomatale, l'évitement de l'ombre, les mécanismes de défense, la communication entre la voie de signalisation à la lumière et la voie de signalisation aux brassinostéroïdes, l'acclimatation au froid, l'élongation racinaire induite par la lumière chez *Arabidopsis*, et le basculement de la phase juvénile à la phase adulte chez le riz (Mao *et al*, 2005 ; Oravecz *et al*, 2006 ; Jang *et al*, 2008 ; Liu *et al*, 2008 ; Yu *et al*, 2008 ; Favory *et al*, 2009 ; Kang *et al*, 2009 ; Crocco *et al*, 2010 ; Jeong *et al*, 2010 ; Luo *et al*, 2010 ; Catala *et al*, 2011 ; Dyachok *et al*, 2011 ; Tanaka *et al*, 2011).

La réponse de croissance des plantes à la lumière implique tout d'abord la perception de la lumière par des photorécepteurs (phytochromes, cryptochromes et phototropines), puis l'activation de multiples et distinctes cascades de transduction du signal pouvant se recouvrir entre elles, pour aboutir à l'activation de gènes de réponse à la lumière par l'action positive de facteurs de transcription. A l'obscurité, AtCOP1 cible les facteurs de transcriptions HY5 (long HYpocotyl 5), LAF1 (Long After Far red light 1) et HFR1 (long Hypocotyl in Far-Red 1) pour ubiquitination et dégradation, conduisant à une répression de la photomorphogénèse

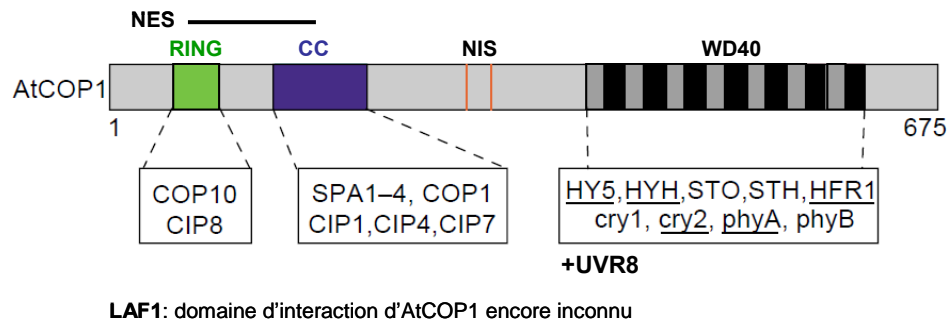


Figure 39 : Domaines protéiques et protéines interactrices d'AtCOP1 (d'après Yi et Deng, 2005 ; Favory *et al*, 2009). AtCOP1 possède un domaine RING en partie N-terminale, un domaine coiled-coil (CC) et 7 répétitions WD40 en partie C-terminale. NES, signal d'export nucléaire ; NIS, signal d'import nucléaire. Les protéines soulignées ont été montrées ubiquitinées par AtCOP1.

(Figure 38, A). En présence de lumière, dans les longueurs d'onde du visible, les photorécepteurs activés répriment la fonction d'AtCOP1, permettant ainsi une accumulation des facteurs de transcriptions activateurs de la réponse de croissance à la lumière et donc le développement photomorphogénique de la plante (Figure 38, B). Un des mécanismes cellulaires répresseur connu de la fonction d'AtCOP1 en présence de lumière est l'exclusion nucléaire de la protéine AtCOP1, empêchant ainsi la dégradation des facteurs de transcriptions par l'E3 ligase. De plus, AtCOP1 empêche la sur-activation du signal de réponse à la lumière par une régulation feedback négative des photorécepteurs à la lumière, en promouvant la dégradation de phyA, phyB et cry2 (Shalitin *et al*, 2002 ; Seo *et al*, 2004 ; Jang *et al*, 2010).

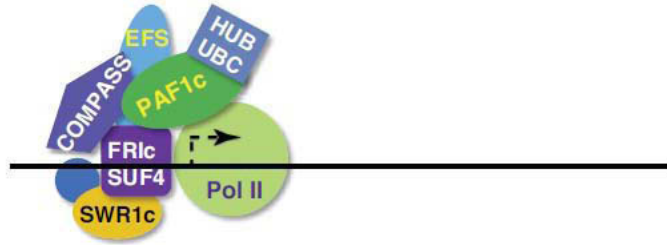
D'autre part, AtCOP1 est un régulateur positif de la signalisation aux UV-B. Suite à une irradiation aux UV-B, AtCOP1 interagit avec UVR8 (UV Resistance 8), promouvant l'expression du facteur de transcription HY5 et conduisant ainsi à l'activation de gènes de réponse aux UV-B (Favory *et al*, 2009 ; Rizzini *et al*, 2011 ; Figure 37, B).

AtCOP1 est une protéine nucléocytoplasmique, constituée de 3 types de domaines : un domaine RING de type C3HC4 en partie N-terminale, un domaine coiled-coil et 7 domaines WD40 en partie C-terminale (Figure 39). Sa distribution intracellulaire est contrôlée par un signal d'import nucléaire situé entre le domaine coiled-coil et les répétitions WD40, et un signal d'export nucléaire en N-terminale de la protéine. Une vingtaine de protéines ont été décrites interagissant avec AtCOP1 (Yi et Deng, 2005 ; Figure 39). Ces protéines sont des cibles ubiquitinées par AtCOP1 mais également des protéines régulatrices de la fonction d'AtCOP1. Pour toutes, excepté LAF1, le domaine d'interaction avec ces protéines dans AtCOP1 a été identifié (Figure 39). Il peut d'ailleurs être remarqué qu'un même domaine peut interagir avec plusieurs protéines différentes. AtCOP1 peut agir sous forme de monomère, d'homodimère ou de complexes.

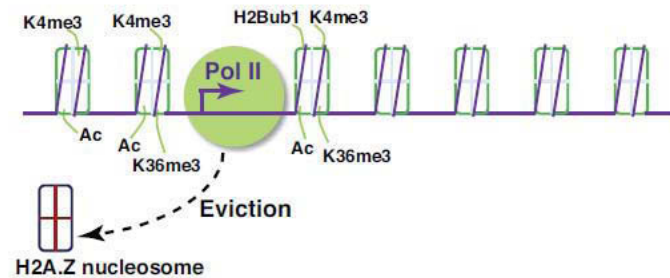
6.2.3.2 La protéine SINAT5

La protéine SINAT5 (SINA of *Arabidopsis thaliana* 5) intervient dans le développement des racines latérales en réponse à l'auxine (Xie *et al*, 2002). En présence d'auxine, le facteur de transcription NAC1 (NAM/NUC-like protein 1) est exprimé dans les primordia racinaires et dans la zone d'élongation de la racine principale, et active la transcription de gènes pour la formation de racines latérales. La protéine E3 RING SINAT5 atténue le signal auxine en ciblant et polyubiquitinant le facteur NAC1 qui est ensuite dégradé

A- Recrutement au locus FLC d'un super-complexe remodeleur de la chromatine



B- Initiation de la transcription



C- Elongation de la transcription

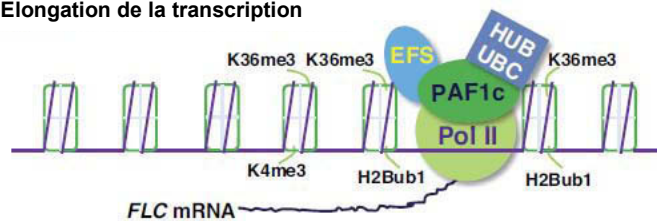


Figure 40 : Modèle d'activation du locus FLC dépendant de FRIGIDA chez Arabidopsis (d'après He, 2012). (A) Le complexe HUB1;2/UBC1;2 est recruté par le complexe PAF1c au site d'initiation de la transcription de gènes du locus FLC. (B) Dans la région d'initiation de la transcription, les nucléosomes sont modifiés par des acétylations, des triméthylations de l'histone H3 (K4me3 et K36me3) et une monoubiquitination de l'histone H2B (H2Bub1), remodelant la chromatine et favorisant ainsi la transcription du gène par l'ARN polymérase II (Pol II). (C) Au cours de l'élongation de la transcription, la monoubiquitination de l'histone H2B maintient l'ouverture de la chromatine en catalysant l'apposition de la marque triméthylée K36me3 sur l'histone H3.

par le protéasome 26S, réinitialisant ainsi la cascade de transduction. SINAT5 est une protéine nucléaire et agit sous forme d'homodimère.

6.2.3.3 La protéine HUB1

Le gène *HUB1* est impliqué dans de nombreux processus de développement : il a été décrit chez *Arabidopsis* dans le contrôle du cycle cellulaire au cours du développement racinaire et des jeunes feuilles (Fleury *et al*, 2007), la dormance de la graine (Liu *et al*, 2007b), la floraison, la défense contre des agents pathogènes nécrotrophes, et très récemment dans le cycle circadien (Himanen *et al*, 2012). Contrairement aux deux protéines citées précédemment, la protéine E3 de type RING HUB1 n'étiquète pas des substrats pour les emmener vers la dégradation. HUB1 monoubiquitine l'histone H2B sur la lysine 146, ce qui conduit à un remodelage de la chromatine (Cao *et al*, 2008).

Chez *Arabidopsis*, *HUB1* possède un paralogue, *HUB2*. Cao et ses collaborateurs (2008) ont montré que, lors de la transition florale, les gènes *HUB1* et *HUB2* sont génétiquement liés aux gènes des E2 *UBC1* et *UBC2*. Les auteurs ont montré par double-hybride que les protéines HUB1 et HUB2 étaient capables de former des homodimères ainsi que des hétérodimères entre elles. Cette interaction a été confirmée et localisée dans le noyau par la technique de BiFC. Cao et ses collaborateurs ont également montré par double-hybride que chaque monomère HUB1 et HUB2 étaient capables d'interagir avec les E2 *UBC1* et *UBC2*, et que le domaine RING des E3 HUB1 et HUB2 était nécessaire à cette interaction. Dans le cycle cellulaire, Fleury et ses collaborateurs (2007) ont montré que HUB1 régule un point de contrôle de la transition entre la phase G2 et M.

La monoubiquitination de l'histone H2B (monoUbH2B) a été décrite dans plusieurs études chez *Arabidopsis* comme favorisant l'apposition d'autres marques épigénétiques qui influencent la transcription de gènes par compaction ou décompaction de la chromatine (Figure 40). Dans la transition florale, la voie *HUB1;2/UBC1;2* génère une monoUbH2B dans les régions codantes de certains gènes du locus répresseur de la transition florale *FLC*. Cette modification favorise ensuite l'apposition des marques épigénétiques H3K4 triméthylée et H3K36 diméthylée sur ces gènes ; l'ensemble de ces modifications conduisant au final à une activation de la transcription de gènes *FLC* et par conséquent à une répression de la floraison (Cao *et al*, 2008). Dans le contrôle du cycle circadien, Himanen et ses collaborateurs (2012) ont montré que la monoUbH2B par HUB1 est également associée à l'apposition de la marque H3K4 triméthylée dans les régions codantes de gènes du cycle circadien, favorisant ainsi la transcription de ces gènes. De plus, Zheng et ses collaborateurs (2012) proposent que la

monoUbH2B par HUB1 influence la dormance de la graine à travers l'apposition de la marque H3K9.

Chez la levure, le complexe d'élongation de la transcription FACT (FACilitates Chromatin Transcription complex) reconnaît et retire les histones H2B monoubiquitinés du nucléosome, relâchant ainsi la structure chromatinienne. Chez Arabidopsis, l'effet de la monoUbH2B par HUB1 sur l'élongation de la transcription n'a pu être encore établi, cependant *HUB1* a été retrouvé génétiquement lié à *FACT* dans le développement des branches et des siliques, mais n'est pas génétiquement lié à *FACT* dans la croissance foliaire et dans le contrôle de la floraison (Van Lijsebettens et Grasser, 2010 ; Lolas *et al*, 2010).

6.3 Les E3 U-box

Les U-box, ou PUB (Plant U-box) chez les plantes, représente la famille de E3 découverte le plus récemment. Les E3 de type U-box sont monomériques et sont facilement identifiables par la présence d'un domaine U-box conservé de 70 acides aminés, identifié pour la 1^{ère} fois dans la protéine UFD2 (Ubiquitin fusion degradation protein 2) chez la levure (Hatakeyama et Nakayama, 2003). Le domaine U-box est un domaine RING modifié, auquel il manque les résidus métal-ligands caractéristiques cystéine et histidine chélatant 2 ions zinc qui confèrent la structure arc-boutée du domaine RING. Le maintien de la structure tridimensionnelle du domaine U-box est alors assurée par des interactions électrostatiques comme des liaisons hydrogènes (Vierstra 2009 ; Yee et Goring, 2009). L'activité E3 ligase des protéines à domaine U-box a été démontrée. Les E3 U-box ne forment pas d'intermédiaire E3-Ub : elles catalysent indirectement le transfert de l'ubiquitine sur la protéine cible en permettant le rapprochement dans l'espace de la E2 et du substrat protéique (Jackson *et al*, 2000 ; Hatakeyama *et al*, 2001). Les génomes d'Arabidopsis et du riz comptent respectivement 64 et 77 gènes codant des protéines U-box (Zeng *et al*, 2008 ; Yee et Goring, 2009). Chez les plantes, les E3 de type U-box ont été reportées pour le moment dans l'auto-incompatibilité, la signalisation par les gibbérélines et l'acide abscissique et la réponse aux stress abiotiques et biotiques (Yee et Goring, 2009). Dans le processus d'autoincompatibilité chez les brassicacées par exemple, lorsque la plante est en présence de son propre pollen, la E3 de type U-box ARC1 est activée suite à sa phosphorylation par le récepteur kinase SRK. ARC1 va alors reconnaître et dégrader le facteur de compatibilité Exo70A1 (Stone *et al*, 2003 ; Samuel *et al*, 2009).

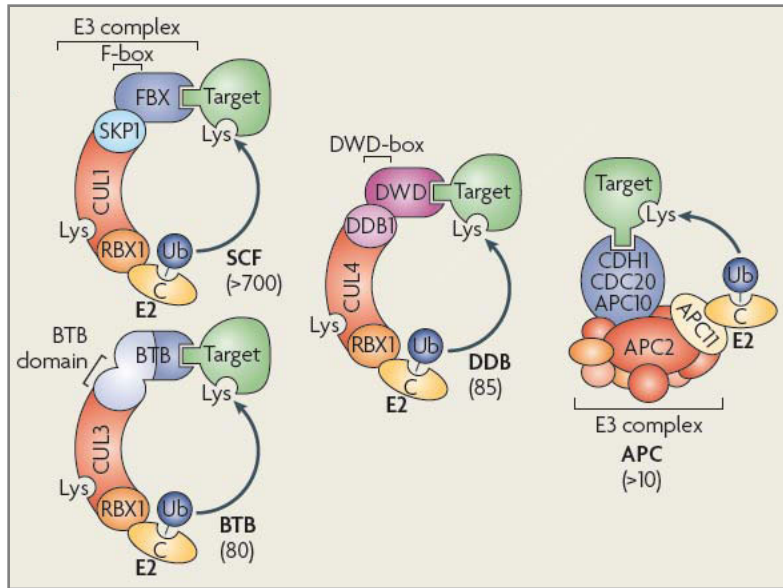


Figure 41 : Les différentes familles de Cullin-RING ligases (CRLs) (d'après Vierstra, 2009). Les nombres entre parenthèses indiquent le nombre de gènes prédits chez *A. thaliana* codant des composants de chaque famille. Les complexes SCF (S phase kinase-associated protein 1(SKP1)-Cullin1- F-box) utilisent la sous-unité F-box pour la spécificité de ciblage de substrat. La sous-unité F-box est reliée à la sous-unité Cullin1 par la sous-unité SKP1 (Suppressor of Kinetochore Protein 1 ou ASK1 chez Arabidopsis). Les complexes CUL4-DDB (Cullin4-DNA Damage-Binding) possèdent une sous-unité DWD pour la reconnaissance du substrat et une sous-unité DDB1 (DNA Damage-binding 1) pour l'interaction avec la sous-unité Cullin 4. Les complexes CUL3-BTB (Cullin 4 - Bric-a-brac-Tramtrack Broad complex) utilisent la sous-unité BTB pour reconnaître le substrat. Cette même sous-unité permet l'interaction avec la sous-unité CUL3. Les complexes APC (Anaphase-Promoting Complex) sont constitués d'au minimum 11 sous-unités, incluant des homologues de RBX1 et de la sous-unité culline (APC11 et APC2 respectivement), ainsi que les sous-unités CDC20 (Cell Division Cycle protein 20), CDH1 (CDC20-homology 1) et APC10, qui peuvent toutes assurer une reconnaissance des substrats protéiques. La sous-unité APC11 permet l'interaction avec l'enzyme E2.

6.4 Les CRLs

Les Cullin-RING ligases sont des E3 multimériques. Elles peuvent être classées en 4 sous-types : les SCF, les CUL4-DDB, les CUL3-BTB et les APC ; en fonction des sous-unités qui composent le complexe (Figure 41). Toutes possèdent les sous-unités culline et RBX1 (RING-box 1), ou des homologues de ces sous-unités dans le cas du complexe APC. La sous-unité RBX1 possède un domaine RING nécessaire pour l'interaction avec l'enzyme E2. Les 4 sous-types de CRL se différencient ensuite par leur module de reconnaissance de la protéine cible.

Les SCFs représentent la plus grande famille de CRL chez les plantes, avec plus de 700 et 687 gènes codant une sous-unité F-box chez *Arabidopsis* et chez le riz respectivement, contre 85 et 78 gènes codant la sous-unité DWD des CUL4-DBD, 80 et 160 gènes codant la sous-unité BTB des CUL3-BTB, et une dizaine de gènes codant le module de reconnaissance du substrat dans les complexes APC (Vierstra, 2009). Pour les complexes SCF, BTB et DDB, l'assemblage du complexe pour l'interaction avec la protéine cible puis le désassemblage du complexe à la fin de la réaction d'ubiquitination sont des processus très dynamiques régulés par les protéines CAND1 (Cullin-Associated and Neddylation-Dissociated 1) et RUB1. La neddylation de la sous-unité Culline, c'est-à-dire l'apposition de la protéine apparentée à l'ubiquitine, RUB1, sur une lysine de la sous-unité Culline favorise l'assemblage du complexe CRL. La déneddylation du complexe, c'est-à-dire le retrait de RUB1 de la sous-unité Culline, est assurée par le complexe COP9 (Constitutive Photomorphogenesis 9)-signalosome. La fixation de la protéine CAND1 sur la sous-unité Culline empêche alors la reformation du complexe (Petroski et Deshaies, 2005 ; Vierstra, 2009).

Tout comme les E3 de type RING et U-box, les CRLs ne forment pas d'intermédiaire avec l'ubiquitine et catalysent le transfert de l'ubiquitine sur le substrat en rapprochant dans l'espace l'enzyme E2 et la protéine ciblée.

D'une manière générale, les CRLs régulent de très nombreux processus chez les plantes. Le complexe APC a été décrit dans la division cellulaire de cellules méristématiques, la différenciation cellulaire, la signalisation à l'auxine et l'endoréplication (Blilou *et al*, 2002 ; Capron *et al*, 2003 ; Vanstraelen *et al*, 2009 ; Mathieu-Rivet *et al*, 2010). Le complexe CUL3-BTB a été décrit dans les mécanismes de défense, dans les voies de signalisation à l'acide gibbérellique, ainsi que dans la signalisation à l'acide abscissique à l'éthylène (Woodger *et al*, 2004 ; Kim *et al*, 2004 ; Christians *et al*, 2009) et très récemment dans la perception de l'acide salicylique (Wu *et al*, 2012). Le complexe CUL4-DDB1 a été décrit

dans la réparation de l'ADN, la prolifération cellulaire, la photomorphogénèse et la signalisation à l'acide abscissique (Chen *et al*, 2006 ; Lee et Zhou, 2007 ; Lee *et al*, 2010). Pour finir, les complexes SCF sont impliqués dans de nombreux processus tels que la signalisation hormonale (auxine, gibbérellines, jasmonate, acide abscissique, éthylène), la photomorphogénèse, le rythme circadien, le développement floral, la sénescence foliaire, le développement des racines latérales, la floraison, l'auto-incompatibilité, et les réponses aux stress biotiques et abiotiques (Xie *et al*, 1998 ; Woo *et al*, 2001 ; Lechner *et al*, 2006 ; Chae *et al*, 2008 ; Vierstra, 2009).

7 Les E3 et le rendement en grain chez les céréales

Comme nous l'avons vu précédemment dans la section 3, le rendement en grains chez les céréales est la résultante de plusieurs composantes mesurables et est fortement impacté par des conditions environnementales défavorables. Nous venons de voir également que les plantes exploitaient fortement l'ubiquitination pour moduler le protéome afin d'assurer correctement des programmes de développement en conditions favorables ou de permettre l'adaptation de la plante à de nouvelles conditions environnementales. Sans surprise, plusieurs exemples de E3 ligases impliquées dans le contrôle de composantes du rendement et dans la réponse de croissance à la sécheresse ont été récemment reportés chez le riz, le maïs et le blé : les protéines E3 ligases de type RING OsSDIR1 (*Oryza sativa* SALT-AND-DROUGHT-INDUCED RING FINGER 1), OsDIS1 (*Oryza sativa* drought-induced SINA protein 1) et TdRF1 (*Triticum durum* RING Finger 1) dans la tolérance à la sécheresse chez le riz et le blé dur (Gao *et al*, 2011 ; Ning 2011a,b ; Guerra *et al*, 2012) ; et les E3 ligases impliquées chez le riz dans la régulation de composantes clés du rendement en conditions favorables comme le complexe APC^{TAD1} (Tillering and dwarf 1) ou APC^{TE} (Tiller enhancer) dans la détermination du nombre de talles ou encore la E3 de type RING GW2 (Grain weight 2) dans la détermination de la largeur et du poids du grain (Song *et al*, 2007 ; Lin *et al*, 2012 ; Xu *et al*, 2012).

7.1 Les protéines E3 RING OsSDIR1, OsDIS1 et TdRF1

La protéine OsSDIR1 est l'orthologue chez le riz de la protéine SDIR1 identifiée chez *Arabidopsis*, impliquée dans la signalisation à l'acide abscissique (ABA) et qui contrôle

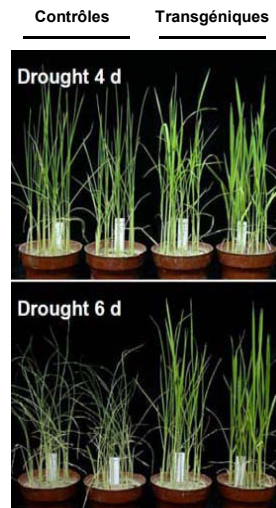


Figure 42 : Phénotypes de lignées de riz transgéniques sur-exprimant *OsSDIR1* comparées à des lignées transformées avec un vecteur contrôle, suite à 4 et 6 jours de stress hydrique (d'après Gao *et al*, 2011).

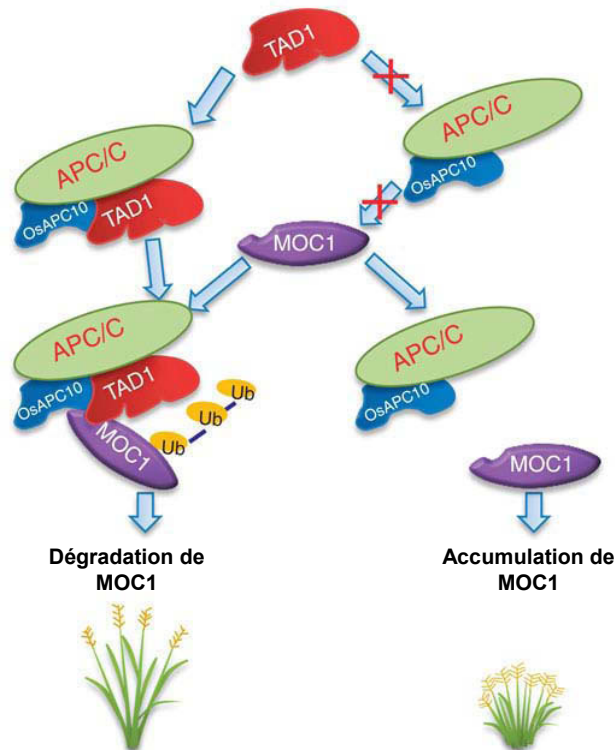


Figure 43 : Modèle de contrôle du tallage chez le riz par le complexe APC^{TAD1} (d'après Xu *et al*, 2012). La sous-unité TAD1 recrute spécifiquement MOC1 en interagissant avec la sous-unité OsAPC10. Le complexe E3 APC^{TAD1} dégrade MOC1, maintenant un taux approprié du régulateur MOC1 et contrôlant ainsi le tallage. En absence de TAD1, MOC1 n'est pas ciblé pour dégradation par le complexe APC, et s'accumule ; augmentant le nombre de talles dans le mutant *tad1*.

l'ouverture stomatale et confère une tolérance des plantes d'*Arabidopsis* à la sécheresse (Zhang *et al*, 2007 ; Gao *et al*, 2011). Des plants de riz sur-exprimant OsSDIR1 présentent une tolérance à la sécheresse accrue (Figure 42). Les auteurs montrent que OsSDIR1 régule l'ouverture stomatale et suggèrent que cette régulation se fait par une voie dépendante de l'ABA, tout comme son homologue chez *Arabidopsis*. Un homologue de OsSDIR1, ZmSDIR1, a également été identifié chez le maïs (Xia *et al*, 2012). Les études de Gao (2011) et Xia (2012) montrent que les OsSDIR1 et ZmSDIR1 sont des protéines membranaires avec un domaine RING de type C3H2C3 et présentent une activité E3 ligase *in vitro*. A l'heure actuelle, la (les) cible(s) de OsSDIR1 et ZmSDIR1 est (sont) encore inconnue(s).

Le gène *OsDIS1* a été identifié comme un régulateur négatif de la tolérance à la sécheresse chez le riz (Ning *et al*, 2011). Des plants de riz sur-exprimant et sous-exprimant *OsDIS1* présentent respectivement une diminution et une augmentation de la tolérance à la sécheresse. La protéine OsDIS1 est une E3 de type RING C3HC4 et présente une activité E3 ligase *in vitro*. OsDIS1 est une protéine nucléocytoplasmique avec une localisation majoritaire dans le noyau. Les auteurs ont montré par double-hybride et par pull-down que OsDIS1 interagit avec la protéine OsNek6 (*Oryza sativa* NIMA-related kinase 6), une sérine/thréonine kinase associée au réseau de tubuline dans le cytoplasme. Les auteurs supposent que OsNek6 est dégradée *in vivo* par OsDIS1 par ubiquitination puisque diverses expériences ont montré que OsNek6 était instable en présence de OsDIS1 *in vivo* et que cette instabilité pouvait être modifiée par un traitement à l'inhibiteur du protéasome, le MG132.

Le gène *TdRF1* a tout d'abord été identifié comme un gène dont l'expression répondait à un stress au froid et à la sécheresse (Mastrangelo *et al*, 2005 ; De Leonardis *et al*, 2007). Récemment, Guerra et ses collaborateurs (2012) ont montré que la sur-expression et la sous-expression de l'homologue de *TdRF1* chez l'orge augmentait et diminuait la tolérance cellulaire à la sécheresse chez l'orge, respectivement. Les auteurs ont également montré que la protéine TdRF1 chez le blé dur interagissait avec le facteur de transcription WBLH1 (Wheat BEL1-type Homeodomain 1), la MAP kinase TdWNK5 (*Triticum durum* WITH NO LYSINE (K)5) et la protéine E3 RING WVIP2 (Wheat Viviparus 1 Interacting Protein 2). TdWNK5 phosphoryle *in vitro* TdRF1. Bien qu'un test d'ubiquitination n'ait pu le mettre en évidence, il est fortement supposé que TdRF1 engendre la dégradation de WBLH1 par la voie UPS. La relation entre les 2 E3 RING TdRF1 et WVIP2 n'est pour l'instant pas établie.

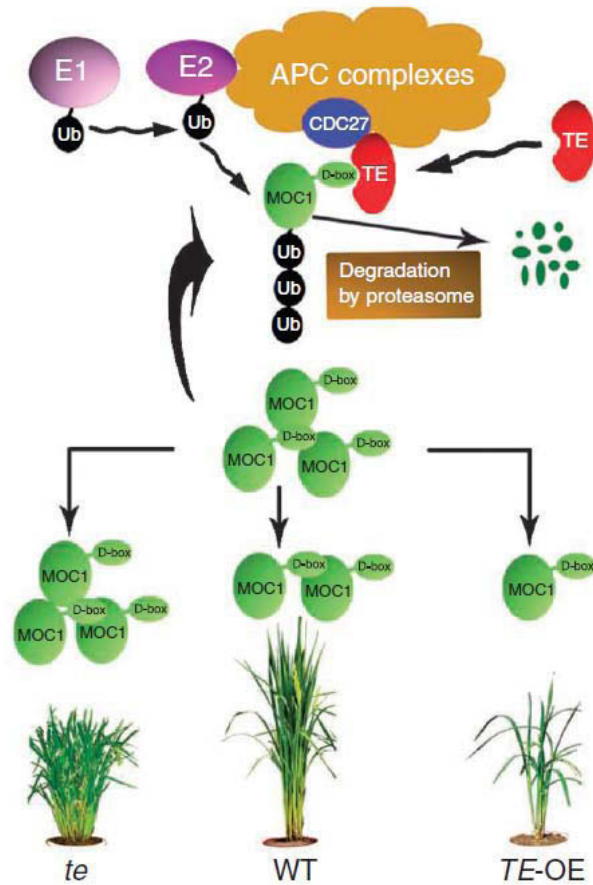


Figure 44 : Modèle de contrôle du tallage chez le riz par le complexe APC^{TE} (d'après Lin *et al*, 2012). La sous-unité TE reconnaît le domaine D-box de MOC1. Le complexe E3 APC^{TE} ubiquitine alors MOC1, conduisant à la dégradation de MOC1 par le protéasome 26S. L'absence de TE permet une accumulation accrue de MOC1 et par conséquent un nombre de talles plus important, dans le mutant *te* par rapport au sauvage (WT). Une sur-expression de TE (*TE-OE*) induit une forte diminution de la quantité du régulateur MOC1 et par conséquent un nombre de talles réduit, par rapport au sauvage.

7.2 Le complexe APC^{TAD1} ou APC^{TE}

En 2012, deux équipes indépendantes reportèrent en même temps la dégradation de la protéine MOC1 par une sous-unité d'un complexe APC (Lin *et al*, 2012 ; Xu *et al*, 2012). Xu et ses collaborateurs (2012) appelèrent cette sous-unité TAD1, du nom du mutant *tad1 tillering and dwarf 1* qu'ils ont identifié, et Lin et ses collaborateurs (2012) la sous-unité TE, du nom du mutant *te tiller enhancer*. Cette sous-unité est un homologue de Cdh1 chez la levure et constitue un adaptateur dans le complexe APC pour la reconnaissance de protéines cibles.

Dans leur étude, Xu et ses collaborateurs montrent que TAD1 reconnaît et ubiquitine spécifiquement le facteur de transcription MOC1, le dégradant et contrôlant par conséquent négativement le nombre de talles chez le riz (Figure 43). Dans le cas d'une perte de fonction TAD1 comme dans le mutant *tad1*, MOC1 ne peut pas être reconnu et dégradé et s'accumule donc, ce qui augmente le nombre de talles. Chez le riz, la majorité des cellules des bourgeons dormants est bloquée en phase G1 du cycle cellulaire. Les auteurs suggèrent que TAD1 recrute MOC1 principalement au cours de la phase G1 et dans les bourgeons de talles. Dans les cellules des bourgeons du plateau de tallage, TAD1 est activé en phase G1 et maintient MOC1 à un faible niveau. Lorsque les cellules des talles basculent en phase S du cycle cellulaire, le niveau de TAD1 chute permettant une accumulation de MOC1 dans la cellule et la production de talles. Les auteurs ont également mis en évidence que TAD1 reconnaît le domaine D-box (destruction box) dans MOC1, et que TAD1 interagit avec une autre sous-unité du complexe APC, OsAPC10, pour cibler MOC1.

Dans leur étude, Lin et ses collaborateurs montrent que TE engendre la dégradation de MOC1 en l'ubiquitinant, diminuant ainsi l'expression du gène d'identité du méristème OSH1 et réprimant donc l'initiation et la formation des méristèmes axillaires (ou bourgeons de talles) (Figure 44). De manière identique, les auteurs suggèrent que TE maintiendrait les cellules méristématiques dans un état mitotique inactif. De plus, Lin et ses collaborateurs mettent en évidence que TE cible MOC1 par la reconnaissance de son domaine D-box, tout comme Xu et ses collaborateurs le montrent, et que TE interagit avec une autre sous-unité du complexe APC, OsCDC27, qui est une sous-unité différente de celle identifiée dans l'étude de Xu. De manière intéressante, Wendt et ses collaborateurs (2001) ont montré que APC10 interagissait avec CDC27 chez l'humain. L'ensemble de ces études laisse donc supposer que le complexe APC constitué par les sous-unités OsAPC10, OsCDC27 et TAD1/TE ubiquitine le facteur de

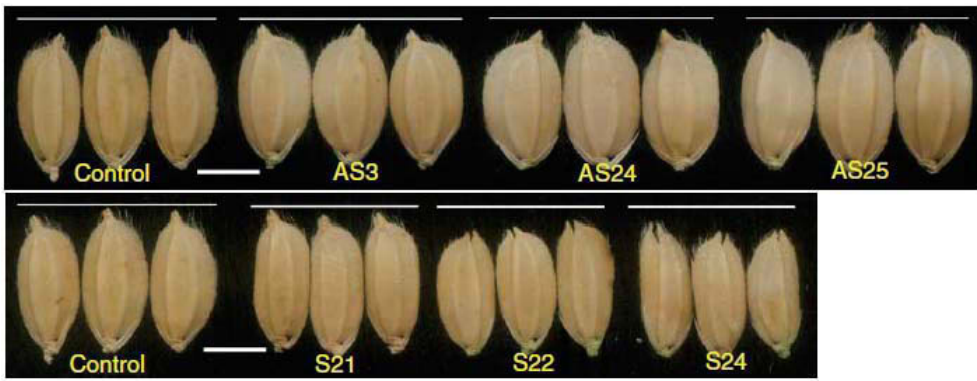
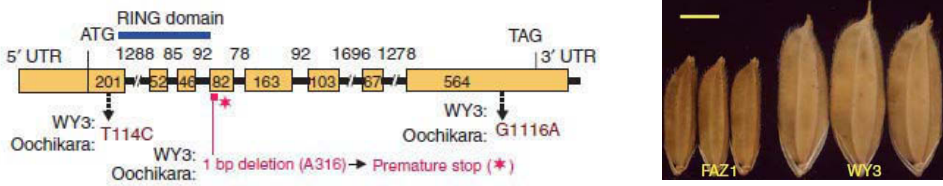
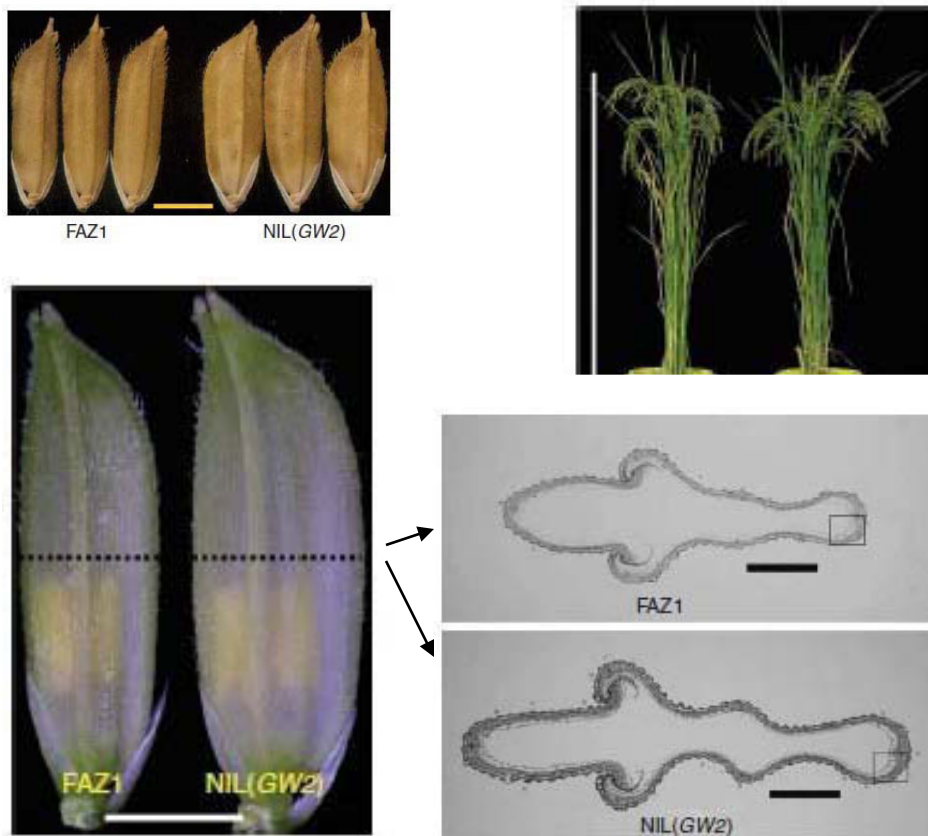
A**B****C**

Figure 45 : *GW2* est un régulateur négatif de la largeur et du poids du grain chez le riz (d'après Song *et al*, 2007). (A) Phénotypes des grains de lignées transgéniques de riz sous-exprimant (AS) ou sur-exprimant (S) *GW2*, comparés aux grains de lignées contrôles. (B) Structure du gène *GW2*, sites des mutations présentes dans les lignées WY3 et Oochikara, et phénotypes des grains des variétés FAZ1 et WY3. La délétion d'une adénine dans l'exon 4 (A316) génère un codon stop prématuré et par conséquent une protéine *GW2* tronquée. (C) Phénotypes des lignées *NIL(GW2)* et de la variété FAZ1. Comparaison des grains (en haut, à gauche), de la hauteur de la plante (en haut, à droite) et des glumelles (en bas, à gauche) entre la lignée isogénique *NIL(GW2)* et la variété FAZ1. En bas, à droite : Sections transversales des glumelles.

transcription MOC1 pour sa dégradation, régulant le nombre de talles et par conséquent impactant le rendement en grain.

8 Le gène *GW2*

GW2 (*grain weight 2*) est un locus à effet quantitatif majeur pour la largeur et le poids du grain de riz (Song *et al*, 2007). Ce locus est situé sur le chromosome 2 du riz, dans une région génomique de 8,2 kb. Une seule ORF prédite a été retrouvée dans ce locus, le gène *GW2*. Ce gène est constitué de 8 exons et 7 introns et son ARN messager code pour une protéine de 425 acides aminés (47 kDa).

Une diminution de l'expression du gène *GW2* par ARN interférence entraîne une augmentation de la largeur du grain de riz de près de 14% tandis que des lignées de sur-expression *GW2* présentent une diminution significative de la largeur du grain de riz (17% ; Figure 45, A). Un variant naturel de cette protéine a été retrouvé notamment dans les variétés de riz WY3 et Oochikara et code pour une protéine *GW2* tronquée. La délétion d'une adénine en position 316 de la séquence nucléotidique de l'ARN messager (exon 4) entraîne un codon stop prématuré et génère une protéine de 115 acides aminés (13 kDa ; Figure 45, B). Les variétés WY3 et Oochikara présentent des grains plus larges que la variété FAZ1 possédant l'allèle sauvage de *GW2* (Figure 45, B). Afin de déterminer la fonction de *GW2* dans le développement du grain de riz, les auteurs ont créé des lignées isogéniques [NIL(*GW2*)] contenant une région de 1,4 cM du locus *GW2* de la variété WY3, dans un fond génétique FAZ1. La lignée NIL(*GW2*) présente des augmentations de 49,8% du poids de 1000-grains, de 26,2% de la largeur du grain, de 10,5% de l'épaisseur du grain et de 6,6% de la longueur du grain de riz, par rapport à la variété sauvage FAZ1 (Figure 45, C). Le rendement en grain par plante, le nombre de panicules par plante et le nombre de jours pour l'épiaison sont augmentés de 19,7%, 27% et de 2,4% respectivement, tandis que le nombre de grains par panicule et la longueur moyenne d'une panicule diminuent de 29,9% et 8%, respectivement dans la lignée NIL(*GW2*). La hauteur de la plante et la largeur de la feuille drapeau ne diffèrent pas significativement entre les 2 lignées (Figure 45, C). Les enveloppes externes du grain (ou glumelles) sont significativement plus larges (+23,4%) dans la lignée NIL(*GW2*) que dans la variété FAZ1 (Figure 45, C). Des coupes transversales des enveloppes externes du grain montrent que la couche de cellules externe du parenchyme est plus longue dans les

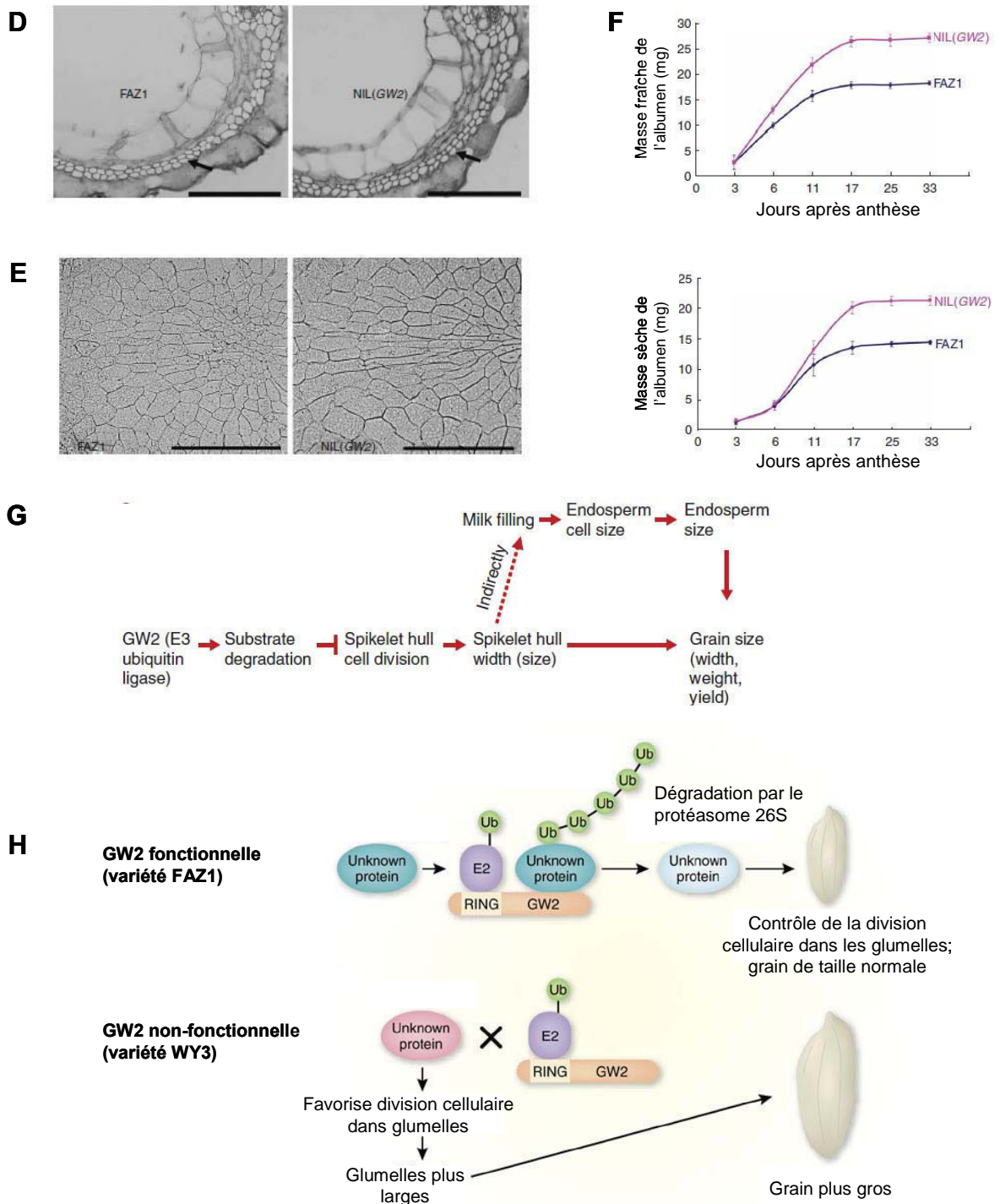


Figure 45 (suite) : **(D)** Comparaison de la couche de cellules externe du parenchyme des glumelles entre la lignée isogénique NIL(*GW2*) et la variété FAZ1. Observations microscopiques de sections transversales des glumelles. **(E)** Comparaison de la structure cellulaire de l'albumen entre NIL(*GW2*) et FAZ1. Observations microscopiques de sections transversales de grain selon l'axe dorso-central. **(F)** Dynamique des masses fraîche et sèche de l'albumen de NIL(*GW2*) et FAZ1, au cours du développement du grain. **(G)** Modèle du rôle de *GW2* dans la régulation de la largeur et du poids du grain, chez le riz. *GW2* cible un régulateur pour dégradation, inhibant la division cellulaire dans les glumelles et influençant alors directement la taille des glumelles, ce qui influencerait indirectement la vitesse de remplissage en réserves du grain, la taille des cellules de l'albumen et par conséquent la taille de l'albumen ; l'ensemble amenant à un contrôle de la taille du grain. **(H)** Modèle de la fonction de *GW2* dans la détermination de la taille du grain, proposé par Matsuoka et Ashikari (2007).

lignées NIL(*GW2*) (+29,6% ; Figure 45, D). Le nombre et la largeur des cellules de cette assise cellulaire sont augmentés de 22,4% et 5,8%, respectivement. Des coupes transversales de grains matures montrent que les cellules de l'albumen sont plus larges (+39,5%) dans la lignée isogénique (Figure 45, E). Aucune différence significative du nombre de cellules de l'albumen n'a été observée entre les 2 lignées. A 17 jours après anthèse, les masses fraîches et sèches de l'albumen sont augmentées de 47,8% et 49,7% respectivement dans la lignée NIL(*GW2*) (Figure 45, F). Bien que les paramètres de la taille du grain de riz soient affectés par *GW2*, les caractères de qualité du grain ne sont pas touchés. En effet, aucune différence significative du contenu en amylose et de la consistance du gel n'a été observée entre les 2 lignées.

Le transcrit *GW2* est exprimé constitutivement dans les parties aériennes et les racines des jeunes plantules de riz, dans les méristèmes inflorescentiels, les jeunes fleurs, les feuilles, les enveloppes externes du grain et l'albumen (à 4 jours après anthèse). Aucune différence d'accumulation du transcrit entre les 2 lignées NIL(*GW2*) et FAZ1 n'a été détectée dans ces différents organes.

La protéine *GW2* possède un domaine RING de type C5HC2, entre les acides aminés en position 62 à 104. A ce jour, ce type de domaine RING n'a été rencontré dans aucune autre protéine. Les auteurs suggèrent que *GW2* représente une nouvelle classe de protéines RING chez les plantes. La protéine *GW2* est une E3 ligase fonctionnelle *in vitro* et présente une localisation subcellulaire cytoplasmique.

En résumé de ces travaux chez le riz, d'après les auteurs, une perte de fonction *GW2* chez le riz entraîne la formation d'enveloppes externes du grain plus larges, ce qui influencerait indirectement la vitesse de remplissage en réserves du grain et la taille des cellules de l'albumen, permettant par conséquent la production de grains plus larges et plus lourds (Figure 45, G et H). Les auteurs suggèrent que *GW2* serait un régulateur négatif de la division cellulaire dans les enveloppes externes du grain de riz. Cette protéine reconnaîtrait spécifiquement une protéine cible pour l'amener à la dégradation par le protéasome 26S. La protéine ciblée par *GW2* au cours de la division cellulaire est encore inconnue à ce jour. Au cours de leur étude, les auteurs ont montré que la protéine *GW2* tronquée dans la variété WY3 possédait toujours l'activité E3 ligase en conditions *in vitro*. La perte de fonction de ce variant de *GW2* serait donc causée par la perte du domaine de liaison du substrat protéique en partie C-terminale de *GW2*.

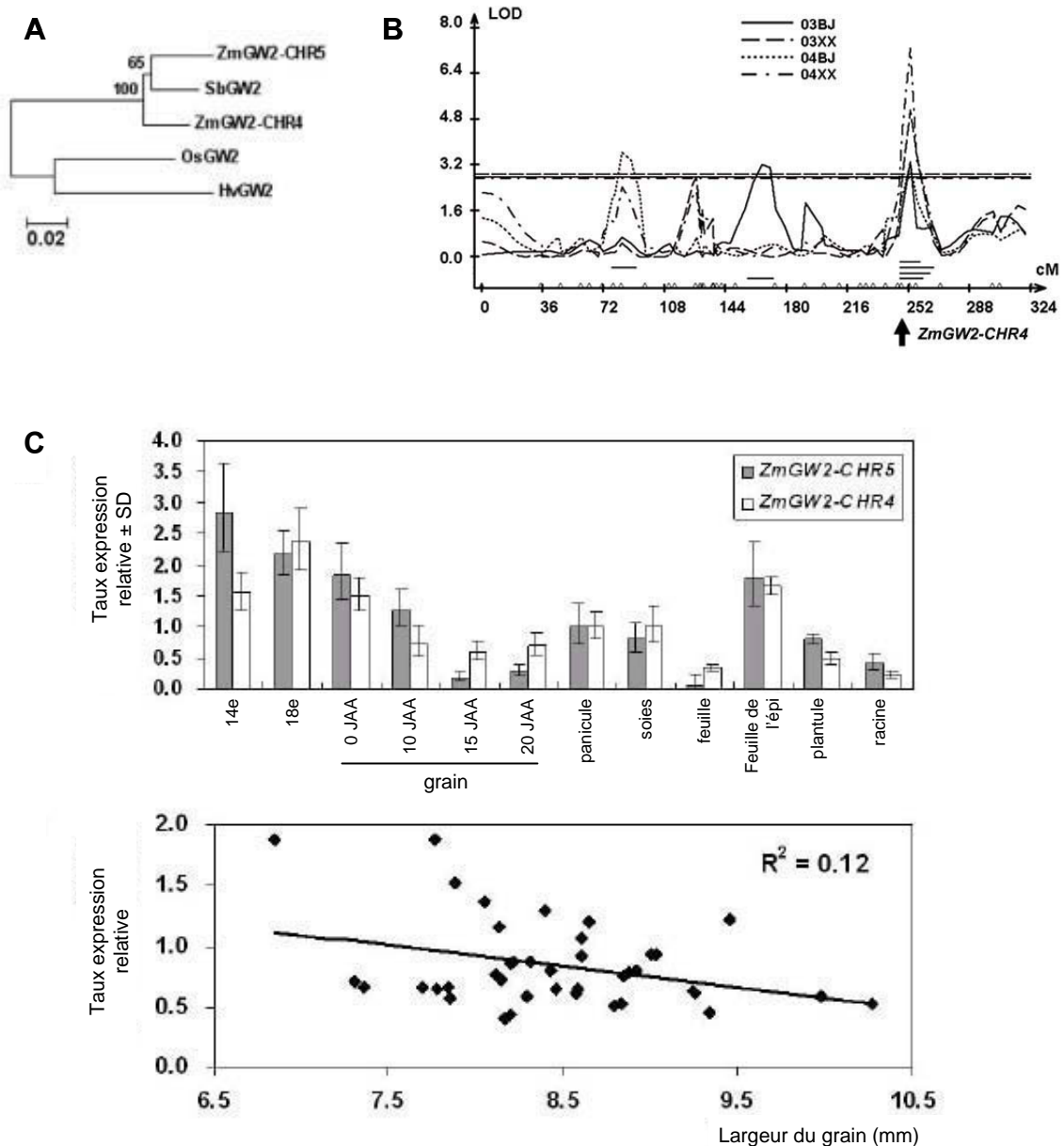


Figure 46 : Les homologues de *GW2* chez le maïs, *ZmGW2-CHR4* et *ZmGW2-CHR5* (d'après Li et al, 2010). (A) Arbre phylogénétique des séquences de *GW2* et de ses homologues chez maïs (*ZmGW2-CHR4* et *ZmGW2-CHR5*), le sorgho (*SbGW2*) et l'orge (*HvGW2*). L'échelle indique le nombre de substitutions en acides aminés par position. (B) Analyse QTL pour le poids de 1000-grains dans une population IF_2 . La position de *ZmGW2-CHR4* est indiquée. 03, population testée en 2003 ; 04, 2004 ; BJ, population testée à Beijing ; XX, Xunxian. (C) Analyse de l'expression de *ZmGW2-CHR4* et *ZmGW2-CHR5*. En haut, profil d'accumulation relative des transcrits *ZmGW2-CHR4* et *ZmGW2-CHR5* dans différents organes de maïs (lignée 87-1). En bas, corrélation entre le taux d'expression relative de *ZmGW2-CHR4* dans des grains âgés de 0 JAA et la largeur des grains (lignée 87-1). 14e et 18e, épis de plantes présentant 14 et 18 feuilles respectivement.

Récemment, afin de déterminer des relations entre 4 gènes impliqués dans le rendement en grains chez le riz, *GS3*, *qSW5/GW5*, *GIF1* et *GW2*, une analyse transcriptionnelle a été menée dans des panicules de plantes transgéniques sous-exprimant ces gènes (Yang *et al*, 2011). Pour mémoire, le gène *GS3* (*grain size 3*) a été identifié dans un QTL de longueur et de poids du grain (Fan *et al*, 2006), et le gène *qSW5/GW5* (*seed width 5*) dans un QTL de largeur et de poids du grain de riz et code pour une protéine nucléaire de type polyubiquitin-binding protein (Weng *et al*, 2008). Le gène *GIF1* (*grain incomplete filling 1*) code pour une invertase de la paroi cellulaire, contrôlant positivement la largeur des grains de riz (Wang *et al*, 2008). L'accumulation du transcrit *GS3* est réduite de 3 fois ($p < 0,001$) et le gène *GIF1* est 2,5 fois plus exprimé ($p < 0,05$) dans des lignées RNAi de *GW2* (Yang *et al*, 2011). La sous-expression de *GW2* chez le riz ne montre aucun effet sur l'accumulation du transcrit *qSW5/GW5*. De même, aucune différence significative de l'expression de *GW2* n'a été détectée dans des plantes sous-exprimant *GS3*, *qSW5/GW5* ou *GIF1*, comparé à des plantes sauvages. Chez le riz, *GW2* régulerait donc positivement l'expression de *GS3* et négativement l'expression de *GIF1*.

8.1 Les gènes homologues de *GW2* chez le maïs, *Zea mays*.

Deux homologues du gène *GW2* ont été identifiés chez le maïs sur les chromosomes 4 et 5, *ZmGW2-CHR4* et *ZmGW2-CHR5* (Li *et al*, 2010a). Ces deux gènes sont constitués de 8 exons et présentent 94% d'identité de séquence en nucléotides entre eux, et 93% d'identité de séquence en nucléotides avec leur homologue chez le riz. Des 2 gènes, *ZmGW2-CHR4* est l'homologue le plus proche de *GW2*, *ZmGW2-CHR5* étant plus proche d'un homologue chez le sorgho (Figure 46, A). Le gène *ZmGW2-CHR4* est situé dans un QTL pour le poids de 1000-grains (Figure 46, B). Une analyse d'association réalisée sur un panel de 121 variétés de maïs dans 3 environnements a permis d'identifier notamment un SNP (transition C/T) dans le promoteur de *ZmGW2-CHR4* associé à la largeur du grain et au poids de 1000-grains dans les 3 environnements ($p < 0,05$). Une insertion-délétion de 8 bases et un SNP (transition C/T) dans le promoteur ainsi qu'un SNP (transition C/T) dans l'exon 8 de *ZmGW2-CHR4* ont également été retrouvés associés à la largeur du grain dans 2 environnements. *ZmGW2-CHR4* et *ZmGW2-CHR5* sont exprimés constitutivement dans les épis, le grain de 0 à 20 jours après anthèse, l'inflorescence mâle, les spathes, les stigmates, les feuilles et les racines. Les 2 gènes présentent un profil d'accumulation en transcrit similaire, avec un coefficient de corrélation

de 0,84 (Figure 46, C). Un maximum d'accumulation de transcrit dans les épis immatures (épis de plants de maïs présentant 14 feuilles) est observé pour les 2 gènes. Bien qu'elle soit très faible ($R^2=0,12$), les auteurs suggèrent qu'il existe une corrélation négative entre le taux d'accumulation de transcrit *ZmGW2-CHR4* et la largeur du grain, indiquant que le gène *ZmGW2-CHR4* serait un régulateur négatif de la largeur du grain et du rendement en grain tout comme *GW2* chez le riz (Figure 46, C). Cependant, les auteurs mentionnent également le fait que malgré la corrélation entre la largeur du grain et le taux d'expression de *ZmGW2-CHR4*, ainsi que l'association entre la largeur du grain et les polymorphismes identifiés dans *ZmGW2-CHR4*, aucun de ces polymorphismes n'affecte le niveau d'expression de *ZmGW2-CHR4*. Le polymorphisme identifié au cours de cette étude dans *ZmGW2-CHR4* comme contribuant à la variation phénotypique est différent du polymorphisme causal dans *GW2* ; en effet la délétion d'une paire de base dans l'exon 4 de *GW2* n'a été retrouvée dans aucune des variétés du panel d'association. Les auteurs concluent leur étude en suggérant que les 2 homologues affecteraient le caractère de largeur du grain par des mécanismes différents chez le riz et le maïs.

8.2 Les homologues de *GW2* chez le blé tendre, *Triticum aestivum*.

Récemment, 3 gènes homologues de *GW2* ont été décrits sur chacun des génomes homéologues A, B et D du blé tendre, cultivar Chinese Spring, sur le groupe de chromosomes 6 (Su *et al*, 2011). D'après les auteurs, les 3 gènes, appelés *TaGW2-A*, *TaGW2-B* et *TaGW2-D*, semblent exprimés puisque 3 séquences différentes correspondant à *GW2* ont été retrouvées dans les banques de courtes séquences exprimées (EST) de blé et dans des produits d'amplification PCR à partir de feuilles et de grains immatures de Chinese Spring. Cependant, aucune des séquences exprimées n'a pu être attribuée par les auteurs à un génome homéologue. Les 3 ADN complémentaires présentent 98% d'identité de séquence en nucléotides entre eux et 87 et 88% d'identité de séquence en nucléotides et en acides aminés respectivement avec *GW2*. Les ARN messagers ont une taille de 1275 pb et codent pour une protéine de 424 acides aminés (47,2 kDa). Le domaine RING de 43 acides aminés en N-terminal des séquences protéiques putatives est strictement conservé par rapport à la séquence protéique de *GW2*. Une analyse transcriptionnelle par RT-PCR montre que *TaGW2*, toutes copies confondues, est exprimé dans les grains de 5 à 25 JAA, dans les feuilles des jeunes plantules, dans la feuille drapeau et dans les jeunes épis (de 10 à 40 mm). Les

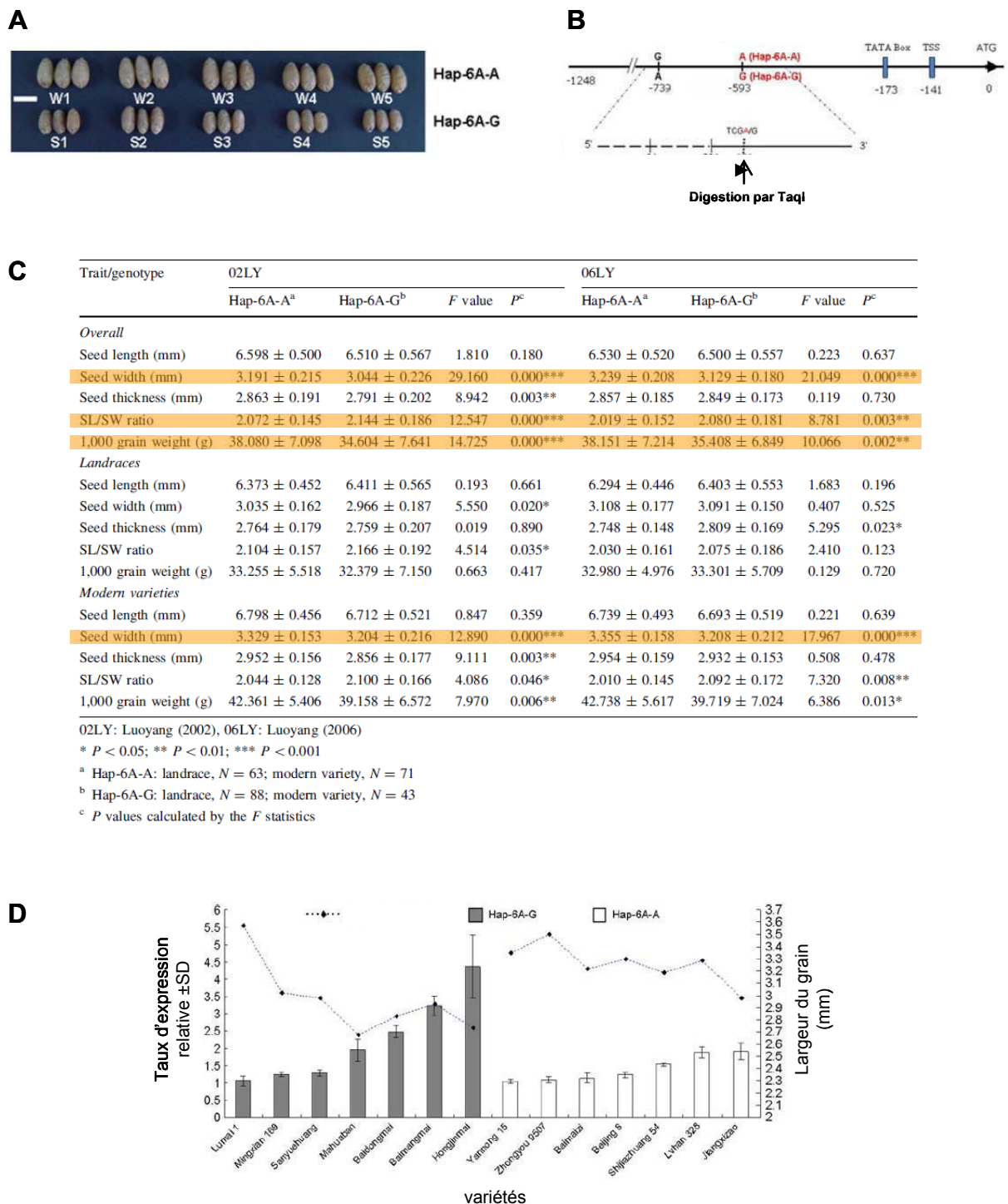
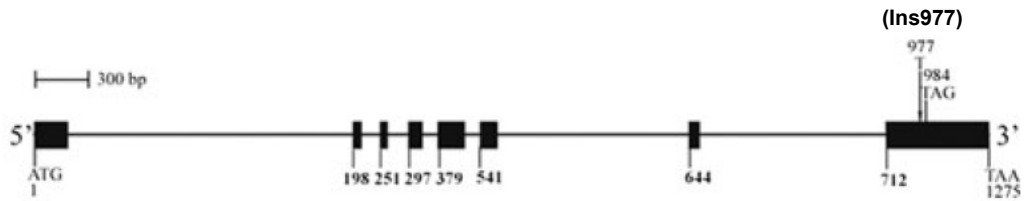


Figure 47 : Un polymorphisme dans le promoteur de *TaGW2-A* est associé avec la taille et au poids du grain dans un panel d'accessions chinoises de blé tendre (d'après Su *et al*, 2011). (A) Comparaison de grains issus de variétés possédant l'allèle A (Hap-6A-A) ou l'allèle G (Hap-6A-G) au marqueur SNP(-593) dans le promoteur de *TaGW2-A*. (B) Positions des marqueurs SNP identifiés dans le promoteur de *TaGW2-A* (en haut). Le SNP en position -593 a permis le développement d'un marqueur CAPS pour un génotypage ultérieur des plantes par digestion enzymatique avec l'enzyme TaqI (en bas). TSS, site d'initiation de la transcription. (C) Analyse d'association entre le marqueur SNP(-593) et les composantes du rendement en grain, sur un panel de 265 accessions chinoises de blé tendre, dans 2 environnements (Luoyang 2002 et 2006). (D) Relation entre le taux d'expression relative de *TaGW2*, toutes copies confondues, et la largeur de grains âgés de 10 JAA, dans 7 variétés possédant l'allèle G et 7 variétés possédant l'allèle A du SNP(-593).

séquences identifiées dans les banques d'EST au début de cette analyse proviennent également de racines, de l'hypocotyle et de l'embryon. Ces résultats suggèrent que *TaGW2*, toutes copies confondues, serait exprimé constitutivement dans le blé tendre. Ces auteurs ont conduit une recherche de marqueurs polymorphes dans les 3 séquences codantes de *TaGW2* dans 10 variétés de blé tendre présentant des phénotypes différents pour la largeur et le poids du grain (5 variétés à gros grains et 5 variétés à petits grains ; Figure 47, A). Cette étude n'a pas permis d'identifier une mutation de délétion entraînant un codon stop prématuré comme dans *GW2*. Cependant, 2 SNP en parfait déséquilibre de liaison ont été identifiés dans le promoteur du gène *TaGW2-A*, des transitions G/A et A/G en position -739 et -593 respectivement, et un marqueur CAPS a été développé sur la base du SNP en position -593 (Figure 47, B). Une analyse d'association a été réalisée pour ce polymorphisme sur un panel de 265 accessions chinoises de blé tendre incluant des variétés modernes et de landraces, dans 2 environnements (1 site, conduit sur les années 2002 et 2006). Le polymorphisme dans le promoteur de *TaGW2-A* a été retrouvé fortement associé à la largeur du grain ($p < 0,001$), au poids de 1000-grains ($p < 0,01$), au ratio longueur/largeur du grain ($p < 0,01$), à la date d'épiaison ($p < 0,001$) et la date de maturité ($p < 0,01$), dans les 2 environnements (Figure 47, C). Une association avec l'épaisseur du grain a été retrouvée pour un seul environnement. Cependant, aucune association n'a été retrouvée pour la longueur du grain. Pour chacune des composantes du rendement présentant une association avec le polymorphisme, l'allèle A est toujours l'allèle favorable. Une analyse d'association pour le polymorphisme dans le promoteur de *TaGW2-A* a également été réalisée dans chacune des sous-populations de landraces et de variétés modernes (Figure 47, C). Dans la sous-population de landraces, une faible association ($p < 0,05$) a été retrouvée entre le polymorphisme étudié et la largeur, l'épaisseur et le ratio longueur/largeur du grain, dans un seul environnement. Dans la sous-population de variétés modernes, le polymorphisme a été retrouvé fortement associé à la largeur du grain ($p < 0,001$), dans les 2 environnements. Une association a également été retrouvée sur les 2 années pour le ratio longueur/largeur du grain ($p < 0,05$ en 2002 et $p < 0,01$ en 2006) et poids de 1000-grains ($p < 0,001$ en 2002 et $p < 0,05$ en 2006), et dans un seul environnement pour l'épaisseur du grain ($p < 0,01$). Une analyse transcriptionnelle de *TaGW2*, toutes copies confondues, par qRT-PCR a été menée sur des grains âgés de 10 JAA, dans 14 variétés issues du panel d'association (7 variétés possédant l'allèle G pour le marqueur SNP en -593 et 7 variétés l'allèle A ; Figure 47, D). A 10 JAA, les auteurs montrent que le taux d'accumulation de transcrit de *TaGW2*, toutes copies confondues, est négativement corrélé à la largeur du grain ($R^2 = 0,36$ et $0,43$ dans les variétés possédant l'allèle G et l'allèle A,

A



B

Trait	Parent		Group mean of F ₂₋₃ families with three different genotypes		
	LK (TT)	CS (tt)	TT	Tt	tt
Overall					
Kernel width (mm)	3.54 ± 0.03	2.82 ± 0.03	3.28 ± 0.01 (B, c)	3.23 ± 0.09 (B, b)	3.10 ± 0.02 (A, a)
Kernel length (mm)	7.94 ± 0.10	6.02 ± 0.05	6.88 ± 0.03 (B, b)	6.86 ± 0.02 (B, b)	6.72 ± 0.03 (A, a)
1,000 kernel weight (g)	57.49 ± 0.88	27.75 ± 0.62	43.00 ± 0.47 (B, b)	41.98 ± 0.28 (B, b)	39.06 ± 0.46 (A, a)
YL					
Kernel width (mm)	3.59 ± 0.02	2.82 ± 0.04	3.28 ± 0.02 (B, c)	3.23 ± 0.01 (B, b)	3.09 ± 0.02 (A, a)
Kernel length (mm)	7.96 ± 0.10	6.01 ± 0.07	6.86 ± 0.04 (B, b)	6.85 ± 0.03 (B, b)	6.67 ± 0.03 (A, a)
1,000 kernel weight (g)	58.60 ± 1.03	28.10 ± 0.86	42.97 ± 0.43 (B, b)	42.07 ± 0.28 (B, b)	39.82 ± 0.45 (A, a)
QS					
Kernel width (mm)	3.55 ± 0.02	2.78 ± 0.04	3.27 ± 0.02 (B, b)	3.22 ± 0.01 (B, b)	3.10 ± 0.02 (A, a)
Kernel length (mm)	7.98 ± 0.10	6.02 ± 0.07	6.93 ± 0.04 (B, b)	6.88 ± 0.03 (B, b)	6.77 ± 0.04 (A, a)
1,000 kernel weight (g)	58.37 ± 1.03	26.73 ± 0.86	42.92 ± 0.61 (B, b)	42.08 ± 0.37 (B, b)	39.15 ± 0.70 (A, a)
QX					
Kernel width (mm)	3.46 ± 0.13	2.89 ± 0.05	3.29 ± 0.01 (B, b)	3.24 ± 0.01 (B, b)	3.11 ± 0.02 (A, a)
Kernel length (mm)	7.86 ± 0.50	6.02 ± 0.24	6.87 ± 0.04 (B, b)	6.86 ± 0.03 (B, b)	6.73 ± 0.04 (A, a)
1,000 kernel weight (g)	54.50 ± 2.60	28.75 ± 1.25	43.08 ± 0.60 (B, b)	41.78 ± 0.36 (B, b)	38.27 ± 0.57 (A, a)

C



Figure 48 : Un polymorphisme dans l'exon 8 de *TaGW2-A* est associé à la largeur du grain dans une population F₂ en ségrégation issue du croisement Lankaodali (LK) x Chinese Spring (CS) (d'après Yang *et al*, 2012). (A) Structure du gène *TaGW2-A* et position du polymorphisme identifié dans l'exon 8 (Ins977). Le Ins977 entraîne un décalage du cadre de lecture et un codon stop prématuré en position 984 (984TAG). (B) Analyse d'association entre les haplotypes au marqueur Ins977 et les composantes du rendement en grain, sur 3 environnements (YL, Yangling ; QS, Qishan ; QX, Qianxian). Test de Duncan; A, B et C indiquent des valeurs significativement différentes à $p \leq 0,001$; a, b et c indiquent des valeurs significativement différentes à $p \leq 0,05$. (C) Comparaison de grains issus de variétés possédant l'allèle TT ou tt au marqueur Ins977. SC, variété Sichuandali ; WM, Wanmai 38 ; MX, Mingxian 169.

respectivement). Les auteurs affirment que le taux moyen d'expression de *TaGW2* est plus important dans les variétés avec l'allèle G que dans les variétés avec l'allèle A, bien qu'un test de Student sur les données fournies dans l'article ne montre pas de différence significative entre les taux d'accumulation de transcrits entre les 2 groupes de variétés ($p=0,12$).

Les auteurs ont également étudié la répartition géographique des allèles dans les diverses variétés chinoises du panel d'association et dans des variétés européennes issues de la core collection mondiale de blé détenue par le centre de ressources génétiques de Clermont-Ferrand (<http://www4.clermont.inra.fr/umr1095>). L'allèle A a majoritairement été sélectionné dans les variétés modernes de blés chinois, tandis que l'allèle G est l'allèle majoritaire dans les variétés de blés européens.

En conclusion de cette étude, les auteurs suggèrent que *TaGW2-A* affecterait négativement la taille finale et le poids du grain par un mécanisme différent de celui de *GW2* chez le riz, à savoir un contrôle de ces composantes du rendement par un niveau d'expression de *TaGW2-A*.

Très récemment, Yang et ses collaborateurs (2012) ont identifié un polymorphisme d'insertion (une base T en position nucléotidique 977) dans l'exon 8 de *TaGW2-A* d'une variété asiatique à large grain, Lankaodali. Cette insertion génère un codon stop prématuré à la position nucléotidique 984 de la séquence codante de *TaGW2-A*, entraînant la production d'une protéine tronquée (328 acides aminés, 37,1 kDa ; Figure 48, A). Ce marqueur a été génotypé sur 372 individus d'une population F2 en ségrégation issue d'un croisement entre la variété à petits grains Chinese Spring et la variété à large grain Lankaodali. Trois génotypes ont été détectés : 75 individus F2 présentent le génotype mutant TT, 184 le génotype hétérozygote Tt, et 68 individus le génotype sauvage tt. Une analyse d'association dans une population F_{2:3} découlant de la population F2, dans 3 environnements (3 sites), a ensuite été réalisée entre ces 3 génotypes et les composantes du rendement suivantes : largeur et longueur du grain, et poids de 1000-grains (Figure 48, B). Les génotypes TT et Tt sont significativement associés ($p<0,001$) à la largeur et longueur du grain, ainsi qu'au poids de 1000-grains, dans les 3 environnements. Le génotype TT augmente en moyenne dans les 3 environnements de 5,8% la largeur du grain, de 2,42% la longueur de grain et de 10,03% le poids de 1000-grains.

Ce polymorphisme a également été retrouvé dans 2 autres variétés asiatiques, Suchuandali et Wanmai 38 (Figure 48, C). La variété Sichuandali présente des grains similaires à la variété Lankaodali, mais avec des valeurs de largeur et longueur du grain plus

grandes, et un poids de 1000-grains plus important. Les grains de la variété Wanmai 38 sont plus larges mais moins longs que les grains de la variété Lankaodali. Aucune information n'est donnée quant au poids de 1000-grains de cette variété.

En conclusion de cette étude et en opposition aux travaux de Su et ses collaborateurs (2011), les auteurs suggèrent que *TaGW2-A* affecterait négativement la taille finale et le poids du grain de blé, par un mécanisme similaire à celui de GW2 chez le riz.

9 Objectifs du travail de thèse

Au sein de l'équipe ABC (Rendement et Adaptation du Blé aux contraintes abiotiques) de l'UMR 1095 GDEC dans laquelle je me suis intégrée, nous nous intéressons entre autres aux mécanismes moléculaires mis en place au cours du développement du grain de blé tendre en conditions optimales et plus particulièrement à la signalisation par la voie UPS. La haute spécificité de ce système de protéolyse ciblée est apportée par l'enzyme clé de la voie UPS, la E3 ligase. Au cours de la dernière décennie, plusieurs E3 ligases ont été reportées chez les céréales dans le contrôle de processus de développement de la plante, impactant d'une manière générale le rendement en grains, dont une plus précisément dans la détermination de la taille finale du grain chez le riz : la E3 ligase de type RING, GW2.

L'objectif principal de ce travail de thèse était d'obtenir l'homologue de *GW2* chez le blé tendre, *T. aestivum*, puis de valider fonctionnellement son rôle au cours du développement du grain. Après l'obtention de connaissances sur les différentes copies de *TaGW2*, la recherche de la fonction de *TaGW2* a été entreprise de deux manières : à l'échelle de la plante entière par validation fonctionnelle en générant des plantes transgéniques de blé tendre sous- ou sur-exprimant le gène *TaGW2* ; et à l'échelle cellulaire par la recherche des partenaires protéiques de *TaGW2* au sein de la cellule, par construction de banques ADNc de grain et leur criblage par la technique de double-hybride.

MATERIEL ET METHODES

Matériel végétal :

1 Les plantes de blé tendre *Triticum aestivum*

Au cours de ce travail de thèse, nous avons majoritairement utilisé la variété Récital. Le choix de cette variété de blé d'hiver a été fait pour le corpus de données transcriptomiques et protéomiques disponibles sur la station pour cette variété. La variété Récital, créée par la société semencière Claude BENOIST, enregistrée au Catalogue Officiel Français des Espèces et Variétés en 1986 et essentiellement cultivée fin des années 80 et début des années 90, fut une variété d'excellente qualité technologique et très appréciée pour sa qualité boulangère. Cette variété est non-barbue, insensible à la photopériode, résistante au froid et à la verse mais sensible à la séptoriose et à la fusariose. Elle est également très sensible à une carence azotée et fut donc choisie, pour cette raison pour 1) l'établissement des populations de cartographie Récital x Renan et Arche x Récital pour la détection de QTLs dans les premiers programmes Génoplante, 2) la production d'EST sur huit stades de développement de grain. L'ensemble de ces ESTs est disponible sur la base de données publiques NCBI DBEST.

Lors des expériences de transgénèse, les variétés Bobwhite et Récital ont été utilisées respectivement pour la production de plantes sur-exprimant *TaGW2-A* et sous-exprimant les homéologues de *TaGW2*.

Pour l'assignation des couples d'amorces homéologues-spécifiques de *TaGW2*, des lignées aneuploïdes de la variété Chinese Spring générées par Sears et ses collaborateurs ont été utilisées (1966). La cartographie des marqueurs *TaGW2_A_Y1061* et *TaGW2_B_Y447* pour les copies A et B de *TaGW2* a été réalisée respectivement sur la population Courtot x Chinese Spring développée par Sourdille *et al* (2003), et la population Renan x Récital développée par Gervais *et al* (2003).

1.1 Production des plantes de blé

Les grains de *Triticum aestivum* cv Récital ont été semés dans des godets (10 x 10 cm) de terreau « spécial semis » à raison de 1 grain par godet, puis placés 2 semaines dans une chambre de culture en conditions contrôlées ($19 \pm 1^\circ\text{C}$, 16h de jour/ 8h de nuit, avec une intensité lumineuse de $200 \mu\text{E} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$, et 45 % d'humidité relative). Après 10 semaines de vernalisation ($4 \pm 1^\circ\text{C}$, 16h de jour/ 8h de nuit, avec une intensité lumineuse de $200 \mu\text{E} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$, et 45 % d'humidité relative), les plantules ont été repiquées dans des pots de 3L (ou des pots de 1L pour les plantes transgéniques T2 de l'évènement T111) contenant du terreau « spécial

blé » (Humustar) et transférées en serre, en conditions contrôlées ($17\pm 2^{\circ}\text{C}$ la nuit, $20\pm 2^{\circ}\text{C}$ le jour ; apport d'éclairage pendant 16h à 450 Watts/m^2 ; 45 % d'humidité relative). Pour l'établissement d'une cinétique de développement du grain et des contingences de place, des séries successives de semis ont été effectuées sur plusieurs mois, avec un semis toutes les semaines.

1.2 Mise en place d'une cinétique de développement du grain de blé

L'identité et la date de floraison des premiers épillets de tous les épis de chaque plante ont été notées. Afin de prendre en compte les variations de températures ayant eu lieu au cours du développement, les stades de développement des grains de blé ont été estimés en temps thermique écoulé à partir de l'anthèse, exprimé en $^{\circ}\text{CJours}$, et non en temps calendaire. Les $^{\circ}\text{CJours}$ correspondent à la somme des températures moyennes journalières accumulées par le grain pendant la période considérée ; la température de la serre étant enregistrée toutes les 10 min. Des grains entiers ont été récoltés aux stades de développement suivants : 40, 80, 120, 160, 180, 220, 240, 260, 300, 400, 500, 600, 700 et 800°CJours ; puis immédiatement plongés dans l'azote liquide et stockés à -80°C . Pour un même stade de développement, seuls les grains provenant d'épillets possédant une date de floraison identique ont été récoltés. Pour chaque stade de développement, 3 réplicats biologiques ont été constitués en regroupant les grains de 4 plantes pour chaque réplicat.

1.3 Obtention d'embryons immatures pour la transgénèse

Ce travail a été réalisé par le personnel de la plateforme « transgénèse blé » de l'UMR. Après 8 semaines de vernalisation pour la variété Récital et aucune période de vernalisation pour la variété Bobwhite, les plantules de *T. aestivum* ont été placées en serre transgénique en conditions contrôlées ($18\pm 2^{\circ}\text{C}$ la nuit, $22\pm 2^{\circ}\text{C}$ le jour ; 16h de jour/ 8h de nuit, intensité lumineuse de $200\mu\text{E.m}^{-2}.\text{s}^{-1}$). Les grains immatures ont été récoltés à 12-14 jours après anthèse et stérilisés ; et les embryons immatures ont été isolés manuellement par dissection puis placés sur un milieu plasmolysant, à l'obscurité.

1.4 La core-collection mondiale de blé tendre.

Pour les analyses d'association, nous avons utilisé l'ensemble des 372 accessions constituant la core collection mondiale de blé tendre, établie par Balfourier et ses collaborateurs (2007). Cette core collection est détenue par le Centre de Ressources Génétiques (CRG) des céréales à paille de Clermont-Ferrand

(<http://www4.clermont.inra.fr/umr1095>). Une core collection est un échantillonnage restreint de variétés représentant le maximum de diversité génétique d'une plus grande population. La structure de cette core collection a été établie par Horvath *et al* (2009), en utilisant le programme STRUCTURE (Pritchard *et al*, 2000) avec 578 marqueurs DArT répartis uniformément sur le génome. La core collection est structurée en 5 groupes géographiques ancestraux (Nord-Ouest de l'Europe, Sud-Est de l'Europe, le groupe CIMMYT-ICARDA, l'Asie et le Népal). Des caractères ont été phénotypés pour l'ensemble des accessions de la core collection par Bordes et ses collaborateurs (2008). Le contenu en protéines du grain (% prot) a été mesuré par spectrométrie de réflectance dans le proche infra-rouge (NIRS). La quantité d'azote par grain a été estimée en multipliant le % prot par le poids de 1000-grains/1000. L'ADN génomique des 13 variétés issues de la core collection utilisées pour la recherche de polymorphisme des copies *TaGW2* a été extrait à partir de feuilles de groupes de 6 plantules par accession considérée, âgées de 3 semaines, comme décrit par Tixier *et al* (1998). Toutes les graines de blé tendre utilisées dans cette étude proviennent du CRG de Clermont-Ferrand, et sont issues d'un seul épi auto-pollinisé. Toutes les plantules d'une même accession peuvent donc être considérées génétiquement identiques.

2 Le tabac (*Nicotiana tabacum*)

2.1 Les plantes de tabac

Une trentaine de graines de *N. tabacum* sont semées dans des pots de 0,5L contenant du terreau « spécial semis », puis placées pendant 2 semaines dans une serre en conditions contrôlées ($17\pm 2^{\circ}\text{C}$ la nuit, $22\pm 2^{\circ}\text{C}$ le jour ; apport d'éclairage pendant 16h à 450 Watts/m^2 ; 45 % d'humidité relative). Les plantules sont alors repiquées individuellement dans des godets de 10 x 10 cm et replacées en serre pendant environ 1 mois. Lorsque les plantes présentent 5 feuilles, elles sont transformées par agroinfiltration pour les expériences d'expression transitoire de protéines chimères. Juste avant l'agroinfiltration, les plantes sont acclimatées 2h en chambre de culture en conditions contrôlées ($25\pm 1^{\circ}\text{C}$, avec une intensité lumineuse de $200\mu\text{E.m}^{-2}.\text{s}^{-1}$, et 50 % d'humidité relative).

2.2 Les suspensions cellulaires de *N. tabacum*

Les suspensions cellulaires de *N. tabacum* (cv BY-2, *Bright Yellow* n°2) sont initiées à partir de cals de *N. tabacum* cultivés sur milieu gélosé. Un à 2 grammes de cals âgés de 3 semaines sont placés dans un erlenmeyer de 100 ml contenant 50 ml de milieu BY-2 liquide (annexe 1), sous une agitation orbitale de 125 rpm, à 25°C , à l'obscurité. Au bout de 3-4 jours, 5 ml de

suspension cellulaire sont repiqués en conditions stériles dans 100 ml de milieu BY-2 frais ; puis dans 160 ml de milieu BY-2 frais 7 jours plus tard. Les cellules de tabac sont maintenues en suspension par agitation orbitale (125 rpm), à 25°C, à l'obscurité. Au-delà de 18 jours de culture, les cellules présentent une coloration grise qui reflète une dégénérescence de la suspension.

Les souches bactériennes

Toutes les souches bactériennes sont cultivées sur le milieu nutritif de Luria-Bertani (LB, annexe 1).

1 Escherichia coli

La souche DH5α

Le génotype de la souche DH5α est : F- Φ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17 (rK-, mK+) phoA supE44 λ- thi-1 gyrA96 relA1.

La souche DH5α présente une efficacité de transformation importante ($>1 \times 10^8$ cfu/μg d'ADN plasmidique) et permet le criblage de transformants par α-complémentation de la β-galactosidase (ou « crible blanc-bleu », lacZΔM15). La souche DH5α amplifie fortement les plasmides, empêche toute recombinaison de l'ADN plasmidique avec l'ADN hôte (recA1) et réduit les clivages non-spécifiques de l'ADN plasmidique lors des étapes de préparation d'ADN plasmidique (endA1 et hsdR17). Enfin, cette souche est parfaitement adaptée à la transformation et l'entretien de plasmides de grande taille.

La souche DH5α utilisée au cours de ce travail de thèse a été achetée chimiquement compétente auprès d'Invitrogen et ne nécessite pas de mise en compétence avant transformation avec du matériel ADN. Un protocole « laboratoire » de mise en compétence de cette souche est tout de même fourni en annexe 2.

2 Agrobacterium tumefaciens

Deux souches d'*Agrobacterium tumefaciens* ont été utilisées : la souche GV3101 (Flückiger *et al*, 2003) pour la transformation transitoire de feuilles de tabac *Nicotiana tabacum* et la souche LBA4404 (Gomord *et al*, 1996) pour la transformation stable de cellules de tabac BY-2. Ces deux souches ont été utilisées pour l'expression de protéines chimères fusionnées à la YFP (Yellow Fluorescent Protein), dans le tabac. Elles portent toutes les deux

Tableau 5 : Génotypes des souches de *S. cerevisiae* utilisées au cours de ce travail de thèse.

Souche de levure	Génotype
NMY51	<i>MATa his3D200 trp1-901 leu2-3,112 ade2 LYS2:: (lexAop)₄-HIS3 ura3:: (lexAop)₈-lacZ (lexAop)₈-ADE2 GAL4</i>
Y187	<i>MATα, ura3-52, his3-200, ade2-101, trp1-901, leu2-3, 112, gal4Δ, gal80Δ, met⁻, URA3 : : GAL1UAS-Gal1TATA-LacZ, MEL1</i>
Y2HGold	<i>MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2 : : GAL1_{UAS}-Gal1_{TATA}-His3, GAL2_{UAS}-Gal2_{TATA}-Ade2, URA3 : : MEL1_{UAS}-Mel1_{TATA}, AUR1-C MEL1</i>
MaV203	<i>MATa; leu2-3,112; trp1-901; his3D200; ade2-101; cyh2^R; can1^R; gal4D; gal80D; GAL1::lacZ; HIS3_{UASGAL1}::HIS3@LYS2; SPAL10_{UASGAL1}::URA3</i>

un plasmide Ti qui code pour les gènes de virulence permettant le transfert et l'intégration de l'ADN-T dans le génome végétal. Ces souches d'agrobactéries sont cultivées sur un milieu LB supplémenté en gentamycine (10µg/ml) pour la souche GV3101, ou supplémenté en streptomycine (500µg/ml) pour la souche LBA4404. Ces gènes de résistance sont apportés par le plasmide Ti.

Les souches de levures *Saccharomyces cerevisiae*

Toutes ces souches de levures sont cultivées sur un milieu nutritif YPDA (Annexe 1). Les génotypes des souches de levures sont mentionnés en Tableau 5.

1 La souche NMY51

La souche de levure NMY51 est commercialisée par Dualsystems Biotech. Cette souche a été utilisée au cours de ce travail de thèse pour les expériences de double-hybride cytoplasmique dans le système DUALhunter (Dualsystems Biotech). C'est une souche haploïde, de type sexuel MATa; auxotrophe pour l'adénine, l'histidine, la leucine et le tryptophane ; et qui possède 3 gènes rapporteurs sous le contrôle de promoteurs répondant au facteur de transcription LexA, uniquement exprimés lors d'interactions protéine-protéine :

- Le gène *HIS3*, qui permet à la levure de synthétiser de l'histidine et donc de pousser sur un milieu minimal sans histidine.
- Le gène *ADE2*, qui permet à la levure de synthétiser de l'adénine et donc de pousser sur un milieu minimal sans adénine.
- Le gène *LacZ* qui code pour l'enzyme β -galactosidase, capable de cliver le substrat chromogénique X-gal lorsqu'il est présent dans le milieu, produisant une coloration bleue de la colonie.

2 La souche Y187

La souche de levure Y2HGold est commercialisée par Clontech. C'est une souche haploïde, de type sexuel MAT α , auxotrophe pour l'adénine, l'histidine, la leucine et le tryptophane ; et délétée pour les gènes *GAL4* et *GAL80*. Cette souche a été utilisée au cours de ce travail de thèse pour la construction des banques d'ADNc de grain de blé tendre dans le système Matchmaker (Clontech).

3 La souche Y2HGold

La souche de levure Y2HGold est commercialisée par Clontech. Cette souche a été utilisée au cours de ce travail de thèse seule et en partenaire de mating de la souche Y187, dans les expériences de double-hybride nucléaire dans le système Matchmaker (Clontech). C'est une souche haploïde, de type sexuel MATa ; auxotrophe pour l'adénine, l'histidine, la leucine et le tryptophane ; délétée pour les gènes *GAL4* et *GAL80*, et qui possède 4 gènes rapporteurs sous le contrôle de promoteurs répondant au facteur de transcription GAL4, uniquement exprimés lors d'interactions protéine-protéine :

- Le gène *AURI-C*, l'allèle mutant dominant du gène *AURI* qui code pour l'enzyme inositol phosphoryl céramide synthase. Son expression dans la levure confère à la levure une forte résistance à l'antifongique Auréobasidine A, qui est un inhibiteur de la synthèse des céramides.
- Le gène *HIS3*, qui permet à la levure de synthétiser de l'histidine et donc de pousser sur un milieu minimal sans histidine.
- Le gène *ADE2*, qui permet à la levure de synthétiser de l'adénine et donc de pousser sur un milieu minimal sans adénine.
- Le gène *MEL1*, qui code pour l'enzyme α -galactosidase capable de cliver le substrat chromogénique X- α -Gal lorsqu'il est présent dans le milieu, donnant une coloration bleue à la colonie.

4 La souche MaV203

La souche de levure MaV203 a été développée par Marc Vidal et est commercialisée par Invitrogen. Cette souche a été utilisée au cours de ce travail de thèse pour les expériences de transactivation de protéine dans le système ProQuest (Invitrogen). C'est une souche haploïde, de type sexuel MAT α ; auxotrophe pour l'uracile, l'histidine, la leucine et le tryptophane ; délétée pour les gènes *GAL4* et *GAL80*, et qui possède 3 gènes rapporteurs sous le contrôle de promoteurs répondant au facteur de transcription GAL4, uniquement exprimés lors d'interactions protéine-protéine :

- Le gène *HIS3*, qui permet à la levure de synthétiser de l'histidine et donc de pousser sur un milieu minimal sans histidine.
- Le gène *URA3*, qui permet à la levure de synthétiser de l'uracile et donc de pousser sur un milieu minimal sans uracile.

- Le gène *LacZ* qui code pour l'enzyme β -galactosidase, capable de cliver le substrat chromogénique X-gal lorsqu'il est présent dans le milieu, produisant une coloration bleue de la colonie.

La souche MaV203 possède également les allèles récessifs de résistance à la cycloheximide (*cyh2^R*) et à la canavanine (*can^R*) pour la ségrégation des plasmides appâts.

Les vecteurs plasmidiques

Les cartes de ces vecteurs plasmidiques sont données en annexe 3.

1 Le plasmide pGEM-T easy

Le plasmide pGEM-T easy, commercialisé par la société Promega, est un vecteur linéarisé de 3015 pb, utilisé pour le clonage de produits PCR. Ce plasmide haute-copie présente une origine *f1* de réplication dans un système bactérien, un gène de résistance à l'ampicilline (100 μ g/ml) et un site de clonage multiple (polylinker) flanqué par les promoteurs de l'ARN polymérase SP6 et T7. L'avantage de ce plasmide est double : 1) la possibilité d'un clonage TA, chacune des extrémités 3' sortantes du vecteur possédant une thymidine terminale, 2) le polylinker est situé dans le gène *LacZ*, permettant un crible blanc/bleu des transformants. Le gène *LacZ* code pour l'enzyme β -galactosidase, capable de cliver le substrat chromogénique X-gal lorsqu'il est présent dans le milieu, produisant une coloration bleue. L'inactivation de la β -galactosidase par insertion d'un produit PCR au polylinker produit une coloration blanche de la colonie de bactérie transformée.

2 Les plasmides Gateway

Une grande majorité des clonages au cours de ce travail de thèse ont été réalisés à l'aide de la technologie Gateway (Invitrogen), dans les vecteurs mentionnés ci-dessous. Tous les vecteurs Gateway, exceptés les vecteurs pEG et le vecteur pSTARGATE, sont commercialisés par Invitrogen. Tous les vecteurs Gateway possèdent entre les sites de recombinaison le gène de toxicité *ccdB*, pour la sélection ultérieure des vecteurs recombinants.

2.1 Les vecteurs donneurs

Le vecteur pDONR221

Ce plasmide de 4762 pb présente un gène de résistance à la kanamycine (50µg/ml), et des sites d'hybridation du couple d'amorce universelle M13 de part et d'autre des sites attP pour le séquençage de l'insert.

Le vecteur pCR8/GW/TOPO

Contrairement au vecteur pDONR221, le clonage de fragment d'intérêt dans ce plasmide ne se fait pas par recombinaison BP, mais par un clonage de type TA. Un clonage dans ce vecteur d'entrée permet un gain de temps puisqu'il évite la réaction BP, mais présente l'inconvénient de ne pas être directionnel. Ce plasmide haute-copie de 2817 pb possède des sites attL1 et attL2 flanquants le site de clonage, et des sites d'hybridation des couples d'amorces M13 et GW1/GW2 permettant le séquençage de l'insert. Le plasmide pCR8/GW/TOPO possède un gène de résistance à la spectinomycine (100µg/ml), ainsi qu'un promoteur T7 permettant une transcription *in vitro* de l'insert.

2.2 Les vecteurs de destination

Le vecteur pDEST17

Le vecteur plasmidique pDEST17 (6354 pb) permet la fusion d'une protéine d'intérêt en C-terminale et dans le cadre de lecture d'une étiquette « 6 histidines » (6xHIS). Le plasmide possède un gène de résistance à l'ampicilline (100µg/ml) pour sa sélection en bactérie.

Les vecteurs pDEST32 et pDEST22

Les vecteurs pDEST32 et pDEST22 sont utilisés pour la génération de protéines de fusion GAL4 DBD (GAL4 DNA Binding domain) et GAL4 AD (GAL4 Activation Domain) de protéine d'intérêt, respectivement, pour les expériences de double-hybride nucléaire dans le système ProQuest (Invitrogen). Ces plasmides sont à faible-copie.

Le plasmide pDEST32 (12266 pb) possède un gène de résistance à la gentamycine (50µg/ml ; pour la sélection en bactérie) et un gène d'auxotrophie LEU2 (pour la sélection en levure).

Le plasmide pDEST22 (8930 pb) possède un gène de résistance à l'ampicilline (100µg/ml ; pour la sélection en bactérie) et un gène d'auxotrophie TRP1 (pour la sélection en levure).

Les vecteurs pEG101, pEG102 et pEG104

Les vecteurs pEG (pEarleyGate) sont des vecteurs binaires adaptés à la transformation de cellules végétales par *A. tumefaciens*, développés par le laboratoire de Craig Pikaard (<http://sites.bio.indiana.edu/~pikaardlab>) et décrits dans Earley *et al* (2006). Les plasmides pEG possèdent deux origines de réplication, l'une pour *E.coli* et l'autre pour *A.tumefaciens* ; et la région de transfert d'ADN-T par le plasmide Ti d'*A.tumefaciens* bornée par les bordures LB (Left Border) et RB (Right Border) responsables de l'intégration d'ADN exogène chez les plantes. Tous les vecteurs pEG utilisés au cours de ce travail de thèse possèdent un gène de résistance à la kanamycine (50µg/ml). Les vecteurs pEG101 et pEG104 permettent l'expression de protéines d'intérêt en fusion N- et C- terminale, respectivement, de la protéine fluorescente YFP (Yellow Fluorescent Protein). Le vecteur pEG102 permet une fusion de la protéine d'intérêt en N-terminale de la protéine fluorescente CFP (Cyan Fluorescent Protein).

Le vecteur pSTARGATE

Le vecteur pSTARGATE a été développé par le CSIRO (Australie) (<http://www.pi.csiro.au/rnai/vectors.htm>). Ce vecteur est une modification du vecteur pSTARLING, lui-même dérivé du vecteur pHELLSGATE (Wesley *et al*, 2001), pour une utilisation dans les plantes monocotylédones et un clonage par le système de recombinaison Gateway. Ce vecteur contient une cassette dessinée pour la production d'ARN double-brin en épingle à cheveu, constituée de :

- un promoteur constitutif suivi d'un intron, issus du gène de l'ubiquitine du maïs
- deux sites d'insertion de la séquence d'intérêt en orientation sens et antisens, par recombinaison homologue (technologie Gateway)
- un intron du gène de la kinase pyruvate déshydrogénase (PDK) d'arabidopsis, permettant le repliement du transgène exprimé en structure en épingle à cheveu
- un gène de résistance à l'hygromycine (30µg/ml) sous le contrôle d'un promoteur fort 35S, pour une sélection des plantes transformées
- une séquence terminateur de la nopaline synthase (NOS)

Le vecteur utilisé au cours de ce travail de thèse pour la construction de la cassette RNAi-TaGW2 est lui-même une adaptation du vecteur pSTARGATE. Le vecteur d'origine a été modifié pour permettre la construction d'une cassette RNAi et d'une cassette de sélection indépendantes. L'ensemble « promoteur 35S-gène de résistance à l'hygromycine-terminateur nos » a été retiré et remplacé par une séquence terminateur de l'octopine synthase (OCS).

3 Les vecteurs plasmidiques du système DUALHunter (Dualsystems Biotech)

Les plasmides pPR3-N et pDHB1 sont utilisés pour la création de protéines proie et appât, respectivement, dans le système de double-hybride cytoplasmique DUALHunter (Dualsystems Biotech). Ils possèdent les origines de réplication pBS et 2 μ ori, pour leur réplication dans *E.coli* et *S.cerevisae* respectivement.

Le plasmide pPR3-N

Le plasmide pPR3-N (6204 pb) est utilisé pour la création de banque ADNc dans le système DUALHunter (Dualsystems Biotech). Il permet la fusion de fragment d'ADN en C-terminale et dans le cadre de lecture de l'ensemble NubG-HA. Le site de clonage comporte les sites SfiA et SfiB pour un clonage par l'enzyme de restriction SfiI. Ce plasmide faible-copie possède un gène de résistance à l'ampicilline (100 μ g/ml, pour la sélection des bactéries recombinantes) et un gène d'auxotrophie TRP1 (pour la sélection des levures recombinantes).

Le plasmide pDHB1

Le plasmide pDHB1 (8961 bp) permet la fusion de la protéine d'intérêt appât en N-terminale et dans le cadre de lecture de l'ensemble Cub-LexA-VP16 ; et en C-terminale et dans le cadre de lecture de la protéine d'ancrage à la membrane Ost4. Le site d'insertion du fragment d'intérêt comporte les sites SfiA et SfiB pour un clonage par l'enzyme de restriction SfiI. Ce plasmide faible-copie possède un gène de résistance à la kanamycine (50 μ g/ml, pour la sélection en bactérie) et un gène d'auxotrophie LEU2 (pour la sélection des levures recombinantes).

4 Les vecteurs du système Matchmaker (Clontech)

Les plasmides pGADT7/ pGADT7-Rec et pGBKT7 sont utilisés pour la création de protéines proie et appât, respectivement, dans le système de double-hybride nucléaire Matchmaker. Ils possèdent les origines de réplication pUC et 2 μ ori, pour leur réplication dans *E.coli* et *S.cerevisae* respectivement.

Les plasmides pGADT7-Rec/pGADT7

Le plasmide pGADT7-Rec (8000 pb) est adapté pour la création de banque ADNc dans le système Matchmaker (Clontech). Les deux vecteurs permettent la fusion de fragment d'ADN en C-terminale et dans le cadre de lecture du domaine GAL4 AD (GAL4 Activation Domain).

Le clonage de fragment d'ADN dans le plasmide pGADT7-Rec se fait par recombinaison entre sites spécifiques (SMART III et CDS III), tandis que le clonage dans le plasmide pGADT7 (7988 pb) est réalisé par enzymes de restriction puis ligation par l'enzyme T4 DNA ligase. Le clonage dans ces deux vecteurs est directionnel. Ce sont des plasmides faible-copie possédant un gène de résistance à l'ampicilline (100µg/ml, pour la sélection en bactérie), un gène d'auxotrophie LEU2 (pour la sélection en levure) ainsi qu'un promoteur T7 suivi d'un tag HA en amont du MCS pour une transcription/traduction *in vitro* d'une protéine fusion HA-proie. La présence d'un NLS en fusion du GAL4 AD permet la localisation nucléaire de la protéine fusion proie. Le séquençage de l'insert est permis par des sites d'hybridation des amorces T7 et « 3'AD sequencing ».

Le plasmide pGBKT7

Le plasmide pGBKT7 (7300 bp) permet la fusion de la protéine d'intérêt appât en C-terminale et dans le cadre de lecture du domaine GAL4 DBD. Le clonage dans ce plasmide se fait à l'aide d'enzymes de restriction, au MCS. Ce plasmide faible-copie possède un gène de résistance à la kanamycine (50µg/ml, pour la sélection des bactéries) et un gène d'auxotrophie TRP1 (pour la sélection des levures). Un promoteur T7 suivi d'un tag c-myc en amont du MCS permettent la transcription/traduction *in vitro* d'une protéine fusion c-myc-appât. Le séquençage de l'insert est permis par des sites d'hybridation des amorces T7 et « 3'DNA-BD sequencing ».

Méthodes

Les séquences et les températures d'hybridation de tous les couples d'amorces utilisés au cours de ce travail de thèse sont mentionnées en Annexe 4.

1 Extraction d'acides nucléiques

1.1 *Extraction d'ADN génomique*

Les tissus sont préalablement broyés à l'azote liquide aux mortier et pilon. L'ADN génomique (ADNg) est extrait par une méthode au CTAB (adaptée d'après Saghai-Marroof *et al*, 1984). L'ADNg extrait est dosé au nanodrop, puis son intégrité est vérifiée sur gel

d'agarose 1%. La pureté des ADN extraits (rapports 260/280nm et 260/230nm) est évaluée au nanodrop sur un échantillon dilué de l'extraction.

1.2 Extraction d'ADN plasmidique

L'ADN plasmidique est extrait à partir de suspensions bactériennes, par la méthode de lyse alcaline (Sambrook et Russell, 2001), ou à l'aide du kit PureLink Quick Plasmid Miniprep (Invitrogen) selon les recommandations du fournisseur. L'ADN plasmidique extrait est dosé sur gel d'agarose 0,8%. La pureté des ADN extraits (rapports 260/280nm et 260/230nm) est évaluée au nanodrop sur un échantillon dilué de l'extraction plasmidique.

1.3 Extraction d'ARN totaux

Les tissus sont préalablement broyés dans de l'azote liquide aux mortier et pilon. Les ARN totaux sont extraits à partir d'environ 250mg de poudre de broyat, à l'aide d'une méthode au chlorure de lithium (adaptée de Bogorad *et al*, 1983), et traités à la DNase I afin d'éliminer toutes traces d'ADN génomique. Les ARN totaux extraits sont dosés au nanodrop, et leur intégrité est vérifiée sur gel d'agarose 2%. La pureté des ARN totaux extraits (rapports 260/280nm et 260/230nm) est évaluée au nanodrop sur un échantillon dilué de l'extraction.

2 Dessin de couples d'amorces

2.1 Homéologues-spécifiques

L'ADNc pleine-longueur de *TaGW2* a été amplifié à l'aide du couple d'amorce TaGW2 FL CDS, dans des feuilles de plantules *T.aestivum* cv Récital âgées de 3 semaines. Les produits d'amplification ont été clonés dans le vecteur pGEM-T easy et séquencés par la société GATC (France). La séquence ADNc de *GW2* (EF447275.1) a été utilisée pour une analyse blast des EST de blé dans Genbank ; et l'alignement entre les 11 ESTs correspondant à *TaGW2* et les ADNc pleine-longueur de *TaGW2* amplifiés dans cv Récital a été réalisé en utilisant le programme ClustalX. La séquence génomique de *GW2* chez le riz a été utilisée afin de prédire la position des introns dans les ADNc *TaGW2*. Tous les couples d'amorces homéologues-spécifiques ont été dessinés sur la base de l'alignement des séquences ADNc *TaGW2-A* et *-B* et de l'EST CK206178. Les couples d'amorces TaGW2-6A, TaGW2-6B, TaGW2-6D ont été utilisés pour l'assignation chromosomique des copies ; les couples TaGW2-6A, TaGW2-6B2, et TaGW2-6D pour la recherche de marqueurs polymorphes; et les

A

Amorce sens copie-spécifique,
fusionnée à une extension M13 (-)

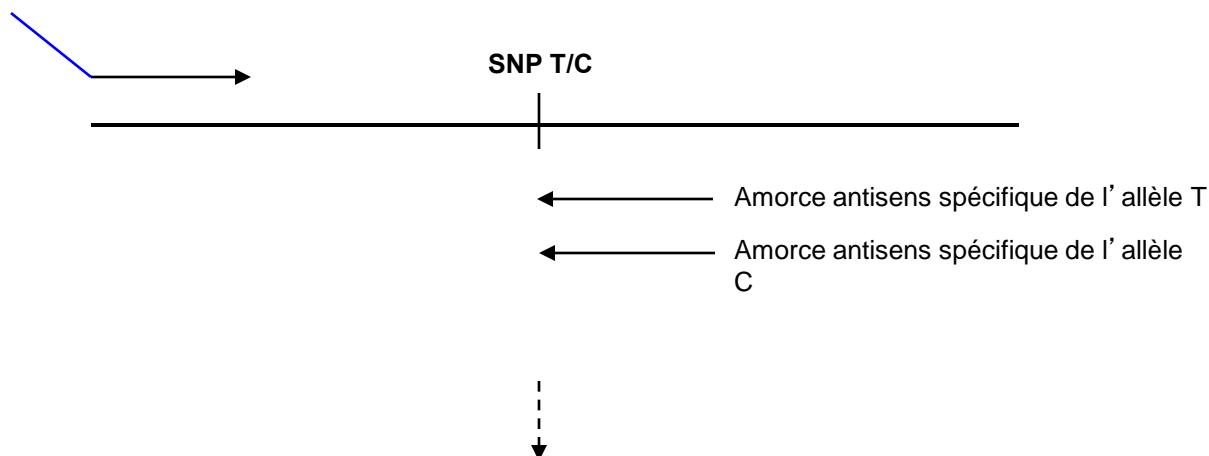
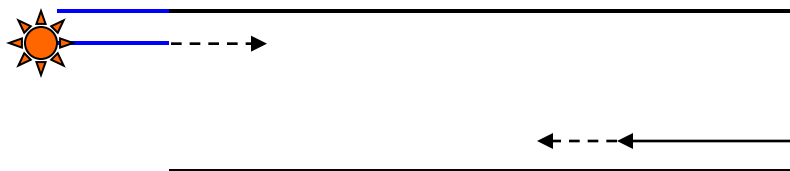
**B**

Figure 49 : Schéma représentatif du principe de génotypage des SNP avec des amorces « extension M13 »: exemple du couple d' amorce A-SNPintron7. (A) L' amorce sens du couple est commune aux deux couples allèle-spécifique, possède une extension M13 (en bleu), et apporte la copie-spécificité d' amplification. L' amorce antisens du couple amène la spécificité vis-à-vis de l' allèle du SNP. (B) Une amorce sens M13 couplée à un fluorophore (étoile orange) permet la détection des produits d' amplification par électrophorèse capillaire.

Tableau 6 : Composition du milieu réactionnel de réaction PCR sur ADN génomique pour le génotypage des SNP. Exemple d' un des deux couples d' amorce allèle spécifique du couple A-SNPintron7.

Composition du mélange réactionnel :

Taq (AmpliTaq Gold - Applied Biosystems)	0,2 U
Amorce sens avec extension M13	50 nM
Amorce antisens	500 nM
Amorce M13 marquée	3,4 pmoles
dNTP	5,5 nmoles

couples TaGW2-A, TaGW2-B et TaGW2-D pour les profils d'expression transcriptionnelle des homéologues.

2.2 Pour le génotypage des marqueurs SNP

Pour chaque copie *TaGW2* considérée, afin d'identifier des marqueurs polymorphes pour cette copie, les séquences des produits d'amplification des 13 variétés de la core collection ont été alignées à l'aide du programme Staden Package (<http://staden.sourceforge.net/>). Des couples d'amorces à la fois homéologue- et allèle-spécifiques pour le marqueur SNP ont été développés (couple « extension M13 »). Pour un SNP considéré, l'une des deux amorces du couple est commune aux deux couples allèle-spécifique et possède une extension M13 (Figure 49). De plus, elle apporte la copie-spécificité d'amplification. L'autre amorce du couple amène la spécificité vis-à-vis de l'allèle du SNP, et cette spécificité est déterminée dans la réaction PCR par le Tm du couple allèle-spécifique.

3 Amplification PCR

Les réactions PCR de routine (PCR semi-quantitative sur ADNc, et sur ADN plasmidique) ont été réalisées à l'aide du kit EconoTaq Plus Green (Lucigen) selon les recommandations du fournisseur. Les amplifications PCR des ADNc pleine longueur des homéologues et pour la recherche initiale de polymorphismes ont été réalisées avec l'enzyme Taq Advantage (Clontech), du fait de son caractère de haute-fidélité, selon les recommandations du fournisseur.

3.1 Génotypage des marqueurs SNP par PCR sur ADN génomique.

Les réactions PCR sur ADN génomique pour le génotypage des marqueurs SNP TaGW2_A_Y1061, TaGW2_B_Y447 et TaGW2_A_R(-593) ont été réalisées comme décrit dans Charmet *et al* (2009), à l'aide des couples d'amorces « extension M13 » A-SNPintron7(T/C), B-SNPintron7(T/G) et A-SNPR(-593)(A/G) respectivement. Brièvement, les réactions PCR ont été réalisées sur 25ng d'ADN génomique, dans le mélange réactionnel mentionné dans le Tableau 6 et selon les programmes PCR mentionnés en Figure 50. Pour un SNP considéré, les réactions PCR sont effectuées sur l'ensemble des individus avec chaque couple d'amorce spécifique de l'allèle, séparément. Au sein du programme PCR, la première phase d'amplification consiste en l'amplification allèle-spécifique, et la deuxième phase en

Couple A-SNPintron7 allèle T			Couple A-SNPintron7 allèle C		
	95°C	5 min		95°C	5 min
x 30	95°C	30 s	1 ^{ère} phase d'amplification : séquence-spécifique	95°C	30 s
	62°C	30 s		61°C	30 s
	72°C	30 s		72°C	30 s
x 8	95°C	30 s	2 ^{ème} phase d'amplification par M13 marquée	95°C	30 s
	58°C	30 s		58°C	30 s
	72°C	30 s		72°C	30 s
	15°C	∞		15°C	∞

Figure 50 : Programmes types d'amplification PCR pour le génotypage des SNP à l'aide d'un jeu de couples d'amorces « extension M13 ». Exemple des couples d'amorce A-SNPintron7 (T/C).

l'amplification des premiers produits d'amplification avec une amorce M13 marquée par un fluorophore, pour une détection ultérieure des produits d'amplification par un séquenceur capillaire. Les produits d'amplification sont ensuite visualisés par électrophorèse capillaire (ABI PRISM®3100, Applied Biosystems) et analysés grâce au logiciel GeneMapper version 3.7 (Applied Biosystems).

3.2 Analyse transcriptionnelle

3.2.1 Transcription inverse

Pour un stade, tissu ou individu considéré, la population d'ARN messagers (ARNm) totaux a été convertie en ADNc simple-brin à l'aide de l'enzyme transcriptase inverse du kit SuperScript II (Invitrogen) et d'oligodT, selon les recommandations du fournisseur. La synthèse d'ADNc simple brin a été réalisée à partir de 2 et 4µg d'ARN totaux pour les profils d'expression de *TaGW2* dans différents tissus et pour les études d'ARN interférence, respectivement. La matrice ARN a été digérée par un traitement à la RNase H (Invitrogen) selon les recommandations du fournisseur, et l'absence de contamination ADN génomique a été vérifiée par PCR sur ADNc à l'aide du couple d'amorce PDI-intron, qui permet l'amplification d'une séquence intronique du gène de la disulfide isomérase chez le blé (*PDI* ; Paolacci *et al*, 2009).

3.2.2 PCR semi-quantitative

Les réactions de PCR semi-quantitative ont été réalisées sur 1µl d'ADNc dilués au 1/10, à l'aide du kit EconoTaq Plus Green (Lucigen) selon les recommandations du fournisseur. Le couple d'amorce TaGW2_318 a été utilisé pour l'amplification du transcrit TaGW2, toutes copies confondues, dans les différents tissus végétatifs de blé tendre.

3.2.3 PCR quantitative

Les réactions de PCR quantitative ont été réalisées à l'aide du kit iQ SYBR Green Supermix (BIORAD), sur l'appareil MyiQ Single-colour Real-Time PCR detection system (BIORAD), selon les recommandations du fournisseur. Les efficacités PCR des couples d'amorces ont été calculées sur des gammes standards de dilution 1/4 (1 :1 à 1 :256) de l'ensemble des populations ADNc de l'étude considérée. Seules les couples d'amorces présentant une efficacité PCR comprise entre 85 et 100% ont été conservés pour les analyses par PCR quantitative. La spécificité des couples d'amorces a été estimée par une analyse de la

courbe de fusion (ou dissociation). Les réactions de RT-PCR quantitative ont été réalisées sur 4µl d'ADNc dilué au 1:40.

Pour les analyses d'expression des homéologues *TaGW2* au cours du développement du grain, trois réplicats biologiques indépendants ont été analysés, avec trois réplicats techniques analysés par réplicat biologique. Les réactions ont été réalisées selon le cycle : 10 min à 95°C ; suivi de 40 cycles à 95°C pendant 15s, et le Tm du couple d'amorces pendant 1min. Afin de s'assurer qu'un seul produit est amplifié par le couple d'amorces, une courbe de dissociation est rajoutée à la fin de chaque réaction de PCR quantitative. La ligne de base de détection et les Ct (« cycle thresholds », cycles seuils) sont estimées de manière automatique à l'aide du logiciel iQ5 software (BIORAD). Les Cts sont exportés et la moyenne des réplicats techniques est calculée. Les taux d'expression relatif bruts sont déterminés en utilisant la méthode du ΔC_t , corrigée pour l'efficacité du couple d'amorces (Pfaffl, 2004). Le gène *Ta2776*, codant une protéine RNase L inhibitory-like, a été décrit par Paolacci et ses collaborateurs (2009) comme étant un des meilleurs et des plus stables gènes de référence chez le blé tendre, pour la normalisation de données d'expression dans divers tissus et dans différents stades de développement de grain. Ce gène de ménage a été testé pour sa stabilité au sein de nos échantillons. Les données d'expression des homéologues *TaGW2* ont été normalisées par rapport à l'expression de *Ta2776*, dans les différents stades de développement du grain de blé. Les valeurs d'expression relatives finales ont été calculées selon la méthode du $2^{-\Delta\Delta C_t}$ (Schmittgen et Livak, 2008), en prenant comme échantillon de référence (ou calibrateur) le taux d'expression relative brut de *TaGW2-A* à 40°CJours.

Dans les lignées RNAi, deux réplicats techniques (extraction ARN) indépendants par individu ont été analysés, avec trois réplicats analysés par réplicat technique. Le gène *Ta54227*, codant une ATPase de la famille AAA (Paolacci et al, 2009) a été testé pour sa stabilité dans nos échantillons expérimentaux et utilisé comme gène de référence. Les valeurs d'expression relatives ont été calculées selon la méthode du $2^{-\Delta\Delta C_t}$, avec la moyenne des valeurs de ΔC_t des nuls-ségrégants de l'évènement de transgénése pris comme échantillon de référence (ou calibrateur interne).

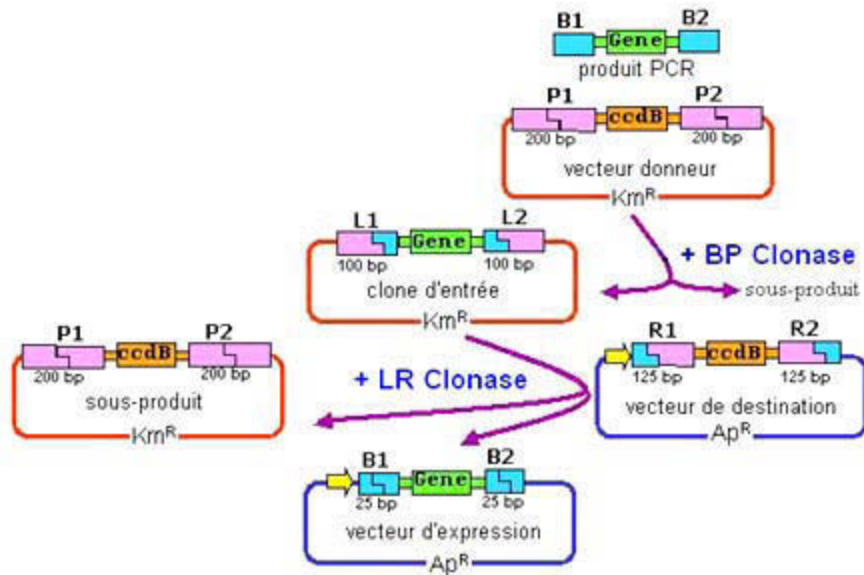


Figure 51 : Principe du clonage d'un gène d'intérêt par recombinaison homologue (Technologie Gateway, Invitrogen). Le clonage par recombinaison *in vitro* se fait en 3 étapes : l'obtention de la séquence d'intérêt à cloner avec à chacune de ses extrémités les sites attB : attB1 et attB2 ; le sous-clonage de la séquence dans un vecteur donneur contenant les sites attP1 et 2 par une recombinaison BP ; et le clonage de la séquence dans un vecteur de destination contenant les sites attL1 et 2 par une recombinaison LR, permettant ainsi l'obtention d'un clone d'expression. Le gène *ccdB* est toxique pour les bactéries *E.coli* et permet de sélectionner uniquement les colonies bactériennes ayant intégré un vecteur recombinant.

4 Clonage dans des vecteurs plasmidiques

Les fragments d'intérêt sont tout d'abord amplifiés par PCR. Un clonage TA ultérieur dans des vecteurs plasmidiques nécessite en particulier à cette étape l'ajout d'un cycle d'élongation finale (7 min, 72°C), afin de rajouter des adénines en 3' des produits d'amplification. Les produits PCR sont séparés par électrophorèse sur gel d'agarose puis purifiés à l'aide du kit NucleoSpin Gel & PCR Clean-Up (Macherey-Nagel), selon les recommandations du fournisseur. Les produits PCR purifiés sont dosés sur gel d'agarose 2%, à l'aide d'une échelle de poids moléculaire de quantification, avant d'être clonés dans les vecteurs plasmidiques par ligation ou par recombinaison. Les méthodes de clonage dans les vecteurs plasmidiques des systèmes DUALHunter et Matchmaker sont données dans les sections x et x, respectivement. Toutes les constructions générées ont été vérifiées par séquençage (Beckman Coulter Genomics).

4.1 *Clonage par ligation*

Le clonage dans le plasmide pCR8/GW/TOPO se fait à l'aide du kit pCR/GW/TOPO TA Cloning kit (Invitrogen), selon les recommandations du fournisseur. La ligation du produit PCR purifié d'intérêt dans le vecteur pGEM-T easy est réalisée par la T4 DNA ligase, à l'aide du kit pGEM®-T Vector System II (Promega), selon les recommandations du fournisseur. Des bactéries *E.coli* souche DH5α chimiquement compétentes (Invitrogen) sont transformées avec le produit de ligation selon les recommandations du fournisseur, et le crible de sélection blanc-bleu pour les transformants pGEM-T est opéré selon les recommandations des fournisseurs.

4.2 *Clonage par recombinaison (Technologie Gateway)*

Les réactions de clonage et de sous-clonage par la technologie Gateway sont réalisées à l'aide des enzymes BP et LR clonases commercialisées par Invitrogen, selon les recommandations du fournisseur. Cette technique de clonage directionnel met en jeu des transferts d'ADN par recombinaison entre sites spécifiques (Figure 51). L'avantage majeur de la technologie Gateway est la rapidité de clonage du fragment d'intérêt, et la possibilité de sous-clonage d'un fragment d'intérêt dans de nombreux vecteurs de destination différents à partir d'un seul et même clone d'entrée. L'ADN d'intérêt est tout d'abord flanqué par des sites de recombinaison attB (attB1 et attB2), par PCR à l'aide de couples d'amorces portant à leur extrémité 5' les sites attB, puis cloné aux sites attP (attP1 et attP2) d'un vecteur donneur

par une réaction BP. Le vecteur recombinant résultant est appelé clone d'entrée, et possède l'ADN d'intérêt flanqué par des sites attL (attL1 et attL2, résultant de la recombinaison entre les sites attB et attP). Le sous-clonage de l'ADN d'intérêt dans un vecteur de destination est alors réalisé par une réaction LR. Les sites attL1 et attL2 du clone d'entrée recombinent respectivement avec les sites attR1 et attR2 du vecteur de destination. Le vecteur recombinant résultant est appelé vecteur recombinant d'expression, et possède l'ADN d'intérêt flanqué des sites attB1 et attB2.

Suite aux réactions de ligation BP puis LR, des bactéries *E.coli* souche DH5 α chimiquement compétentes (Invitrogen) sont transformées avec les produits de ligation selon les recommandations du fournisseur.

5 Mutagenèse dirigée

Les divers mutants de substitutions de ce travail de thèse ont été générés par mutagenèse dirigée à l'aide du kit QuickChange II Site-directed mutagenesis kit (Agilent) selon les recommandations du fournisseur.

6 Transformation de cellules de tabac

6.1 *Mise en compétence d'Agrobacterium tumefaciens*

Une préculture d'*A. tumefaciens* (souche LBA4404 ou GV3101) est diluée dans 150ml de milieu LB. Après 3-4 heures d'incubation en agitation à 28°C (DO₆₀₀ = 0,5), les bactéries en phase exponentielle de croissance sont centrifugées à 3000xg, pendant 20min, à 4°C. Le culot bactérien est lavé dans 10ml de tampon TE (Tris pH8 100mM ; EDTA 1mM) équilibré à 4°C puis resuspendu dans 20ml de milieu LB frais. Des aliquots de 500 μ l sont utilisés directement pour la transformation ou congelés dans l'azote liquide puis stockés à -80°C.

6.2 *Transformation d'Agrobacterium tumefaciens*

La méthode décrite ci-après est une adaptation du protocole décrit dans Höfgen et Willmitzer (1988). Pour la transformation, 500 μ l de bactéries *A.tumefaciens* compétentes sont mélangés avec 1 à 5 μ g d'ADN plasmidique, puis le choc thermique suivant est réalisé : 5 min dans la glace, 5 min dans l'azote liquide et 5 min à 37°C (Bain-marie). La suspension bactérienne est ensuite diluée dans 1 ml de milieu LB puis placée en agitation (180rpm)

durant 3-4 heures à 28°C, pour permettre aux agrobactéries transformées d'exprimer le gène de résistance. Un aliquot de 200µl est alors étalé sur boîte de milieu LB complété en kanamycine (100µg/ml) et en streptomycine (500µg/ml), ou en gentamycine (25µg/ml) selon la souche LBA ou GV, respectivement. Après deux jours d'incubation des boîtes de pétri à 28°C, des colonies isolées sont prélevées et mises en culture liquide pour la transformation de cellules végétales.

6.3 Transformation de suspensions cellulaires de tabac BY-2

Un ml d'une culture végétale âgée de 3 jours est incubé en présence de 50µl d'une préculture d'*A.tumefaciens* LBA4404 transformée, préalablement cultivée pendant 16h à 28°C dans du milieu LB supplémenté en kanamycine et streptomycine. La co-culture est maintenue pendant 2 jours à 25°C, sans agitation, à l'obscurité. Les cellules BY-2 transformées sont centrifugées pour éliminer l'ancien milieu durant 3 min, à 500xg. L'excès d'agrobactéries est ensuite éliminé par 3 lavages successifs avec du milieu BY-2 frais (Annexe 1). Les cellules végétales transformées sont alors étalées sur des boîtes de milieu BY-2 (Annexe 1) supplémentées en céfotaxime (250µg/ml) afin d'éliminer les éventuelles agrobactéries restantes, et en kanamycine (100µg/ml). Après un mois de culture à l'obscurité, à 25°C, des microcals transformés apparaissent. Ils sont prélevés puis transférés sur des boîtes de milieu BY-2 frais supplémentées en antibiotique. Les cals sont observés au microscope confocal.

6.4 Infiltration de l'épiderme inférieur de feuilles de *Nicotiana tabacum*

L'infiltration de l'épiderme inférieur de feuilles de *N.tabacum* par *Agrobacterium* GV3101 est réalisée sur des jeunes plantes de tabac âgées d'environ 5 à 6 semaines, soit présentant en général 5 feuilles (cf section 2.1, Matériel). Les colonies d'agrobactéries transformées sont cultivées dans du LB, à 28°C, jusqu'à la phase stationnaire de croissance, soit environ 20h. Un ml de culture d'agrobactéries est centrifugé à 5000g pendant 3 min, puis le culot est lavé deux fois dans 1 ml de milieu d'infiltration frais (MES 50mM, pH5,6 ; Na₂PO₄ 2mM ; acétosyringone 100mM ; glucose 0,5%). Les agrobactéries sont finalement reprises dans le milieu d'infiltration frais pour une densité optique à 600nm de 0,5. La suspension est placée dans une seringue de 1 ml sans aiguille pour infiltrer les feuilles par simple pression de la seringue sur l'épiderme inférieur des feuilles. Dans le cas d'une co-infiltration des constructions CFP et fusion YFP, des volumes égaux de chacune des suspensions d'agrobactéries transformées sont mélangés et placés dans la seringue. Deux jours

après, les feuilles exprimant les protéines chimères sont observées au microscope confocal (CSLM, confocal laser scanning microscope).

7 Observations en microscopie confocale

La microscopie confocale permet de pratiquer des coupes optiques virtuelles d'un échantillon et de n'enregistrer donc que l'image de la fluorescence émise dans un plan focal choisi. La microscopie confocale offre divers avantages par rapport à la microscopie conventionnelle à épifluorescence, à savoir l'élimination du signal fluorescent provenant d'autres plans, l'observation simultanée de différents fluorochromes et l'acquisition de séries de sections optiques, permettant la reconstruction 3D d'un échantillon ou l'élaboration d'un film. De plus, dans le cas d'un microscope confocal, la source de lumière est un laser donnant des raies de lumière monochromatique de longueur d'onde très précise. Le microscope confocal à balayage utilisé au cours de ce travail de thèse est un microscope LEICA TCS SP5-CSLM, hébergé par la plateforme de microscopie de l'UMR 6293 Génétique Reproduction et Développement (GReD, Clermont-Ferrand). Il est équipé d'un laser Argon multi-raies (458, 476, 488, 496 et 514 nm), d'une diode DSPSS (Diode Pumped Solid State) à 561 nm, et d'un laser Hélium/Néon à 633 nm. Il dispose également d'un répartiteur acoustico-optique AOBs (acoustico-optical beam splitter) qui dirige la lumière d'excitation du laser sur la préparation et sépare de la préparation la lumière fluorescente de retour et la lumière d'excitation réfléchie. Ce système permet l'utilisation simultanément jusqu'à 8 longueurs d'ondes d'excitation.

Au cours de ce travail de thèse, les images CSLM de cellules de tabac co-exprimant à la fois la CFP et une protéine fusion YFP ont été obtenues par excitation simultanée du laser Argon aux longueurs d'ondes 458 nm pour CFP et 514 nm pour YFP. Les signaux de fluorescence ont été séparés à l'aide de l'AOBS, et les émissions de fluorescence de CFP et YFP ont été collectées respectivement par le photomultiplicateur 1 (468 à 502 nm) et 2 (525 à 600 nm). Afin d'éviter les phénomènes de diaphonie (ou « crosstalk ») entre CFP et YFP, le mode d'acquisition séquentiel a été utilisé afin d'acquérir les signaux des canaux les uns après les autres. Les images ont ensuite été analysées par le logiciel LAS AF Lite (LEICA).

Expression de la protéine 6xHis-TaGW2-A en bactérie *E.coli* souche « rescue »:

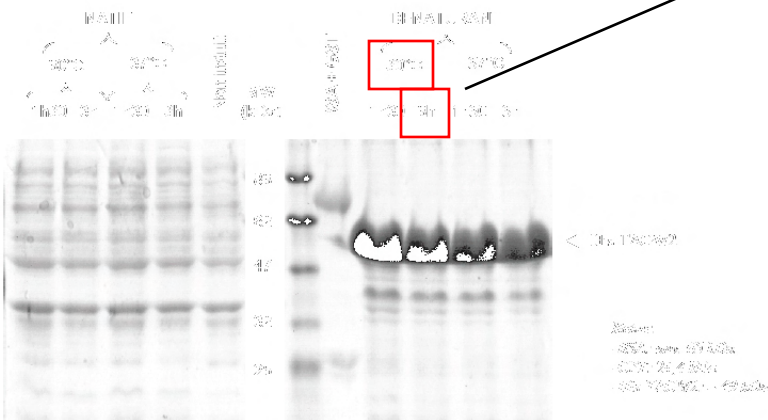
- Test de température de croissance des bactéries pour l'expression de la protéine d'intérêt
- Cinétique de temps après induction de l'expression à l'IPTG (1 mM), pour le prélèvement pour l'extraction protéique



Deux fractions récupérées au cours de l'extraction protéique :

- Lyse des bactéries en tampon natif (PBS) par sonication, centrifugation pour clarification
 - Récupération de la fraction de protéines solubles (+ culot)
- Culot repris dans un tampon PBS/Urée, soniqué et clarification de l'extrait par centrifugation
 - Récupération de la fraction de protéines insolubles

Dépôt sur gel SDS-PAGE des deux fractions



La protéine 6xHis-TaGW2-A est exclusivement présente dans les corps d'inclusion

Purification de la protéine 6xHis-TaGW2-A sur colonnes d'affinité nickel (IMAC):

- Fixation sur colonne dans un tampon PBS/Urée pH7,4
- Elution soit en:
 - PBS/Urée(4M), 150 mM Imidazole (purification en conditions dénaturantes);
 - soit en PBS, 150 mM Imidazole (purification en conditions natives)

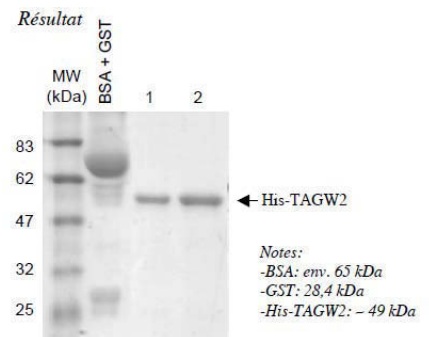


Photo du dépôt sur gel SDS-PAGE de deux fractions d'élution issue de la purification en condition native

Figure 52 : Compte-rendu de la production et de la purification de la protéine 6xHis-TaGW2-A par la société Proteogenix.

8 Test d'autoubiquitination *in vitro*

La protéine fusion 6xHis-TaGW2-A a été produite en bactéries et purifiée par la société Proteogenix (France ; Figure 52). Brièvement, la séquence complète codante de TaGW2-A a été fusionnée en C-terminale et dans le cadre de lecture d'une étiquette 6xHis, dans le vecteur pDEST17. Une production de la protéine fusion 6xHis-TaGW2-A en *E.coli* souche BL21 entraînant une très faible expression de la protéine et son accumulation dans les corps d'inclusion, la protéine fusion a donc été produite dans une souche *E.coli* appelée « rescue » adaptée à l'expression difficile de protéines. La protéine de fusion a été extraite en conditions dénaturantes (PBS/Urée), puis purifiée sur colonnes d'affinité nickel (Figure 52). La protéine de fusion a notamment été éluée dans un tampon final « renaturant » (PBS ; Imidazole 150mM).

Les essais *in vitro* d'ubiquitination de la protéine 6xHis-TaGW2-A ont été réalisés comme décrit dans Song et al (2007), avec quelques modifications : 0,25µg d'enzyme E1 (A1S9, Calbiochem), 0,2µg d'enzyme E2 (GST-UbcH5b, Calbiochem), 5µg d'ubiquitine (GST-Ub, Calbiochem) et 1,2µg de la protéine fusion 6xHis-TaGW2-A purifiée ont été incubés dans un bain-marie à 30°C, pendant 5h, dans un mix réactionnel de 30µl (Tri-HCl 50mM, pH7,4 ; MgCl₂ 2,5mM ; ATP 6,6mM ; et DTT 0,5mM). La réaction a été arrêtée par l'ajout de tampon de charge 5X SDS-PAGE et un chauffage à 95°C, pendant 5 min. Les échantillons protéiques (15µl) ont été séparés sur gel SDS-PAGE 8% ; et les protéines ubiquitinées ont été révélées par immunodétection avec un anticorps de lapin anti-Ub (P4D1, Santa Cruz Technologies).

9 Analyses statistiques

9.1 Cartographie génétique

Les analyses de liaison ont été réalisées à l'aide du logiciel MAPMAKER/EXP 3.0 (Lander *et al*, 1987). Les groupes de liaison ont été obtenus à l'aide de la fonction « group » en utilisant une valeur de LOD de 4 et une fréquence maximale de recombinaison de 0,35. La commande « order » a été utilisée pour déterminer l'ordre des marqueurs ; et la fonction de cartographie Haldane a été utilisée afin de transformer les fréquences de recombinaison en distances en cM sur la carte. La fonction « compare » a été utilisée lorsque l'ordre des marqueurs ne pouvait pas être défini, en prenant en compte l'ordre le plus vraisemblable. Les

fonctions « error-detection » et « genotypes » ont été employées pour déterminer la présence d'éventuelles erreurs de génotypage.

9.2 Analyses d'association

Les analyses d'association entre les allèles aux marqueurs et les variations phénotypiques ont été entreprises avec le logiciel TASSEL 2.0, en utilisant le modèle GLM (General Linear Model). Le marqueur testé a été considéré comme un effet fixé. Le modèle GLM prend en compte la structure de la core collection : la matrice Q a été utilisée comme co-variable afin d'indiquer la contribution de chaque accession à chacun des 5 groupes ancestraux inférés. La significativité des associations entre les marqueurs et les caractères a été évaluée par un test de Fischer.

9.3 Significativité des différences observées dans les lignées transgéniques et dans les analyses transcriptionnelles

Au sein de chaque lignée transgénique, l'effet statistique de la présence de la cassette RNAi sur les caractères finaux du grain a été étudié par une analyse de variance entreprise à l'aide du logiciel SAS/STAT® (procédure GLM, SAS institute 2002). Les comparaisons entre les moyennes ont été réalisées à l'aide d'un test SNK (Student-Newman-Keuls ; seuil 5%).

Pour l'analyse de l'expression des homéologues au cours du développement du grain, une comparaison de l'accumulation des transcrits entre les stades de développement a été entreprise indépendamment pour chaque copie de *TaGW2*. Pour chacune des analyses transcriptionnelles de ce travail de thèse, une analyse de variance a été entreprise ; et les comparaisons entre les moyennes ont été réalisées par des tests de Scheffé ou de Student.

10 Extinction *in planta* de l'expression de *TaGW2* par ARN interférence.

L'ARN interférence est une technique basée sur la dégradation post-transcriptionnelle séquence-spécifique d'ARN messenger, suite à l'introduction d'ARN double-brin dans l'organisme transformé.

L'interférence ARN, historiquement appelée PTGS (Post-transcriptional gene silencing) chez les plantes, a tout d'abord été identifiée chez le pétunia, dans les années 90, par l'équipe de Jorgensen et ses collaborateurs, qui travaillait sur les mécanismes de

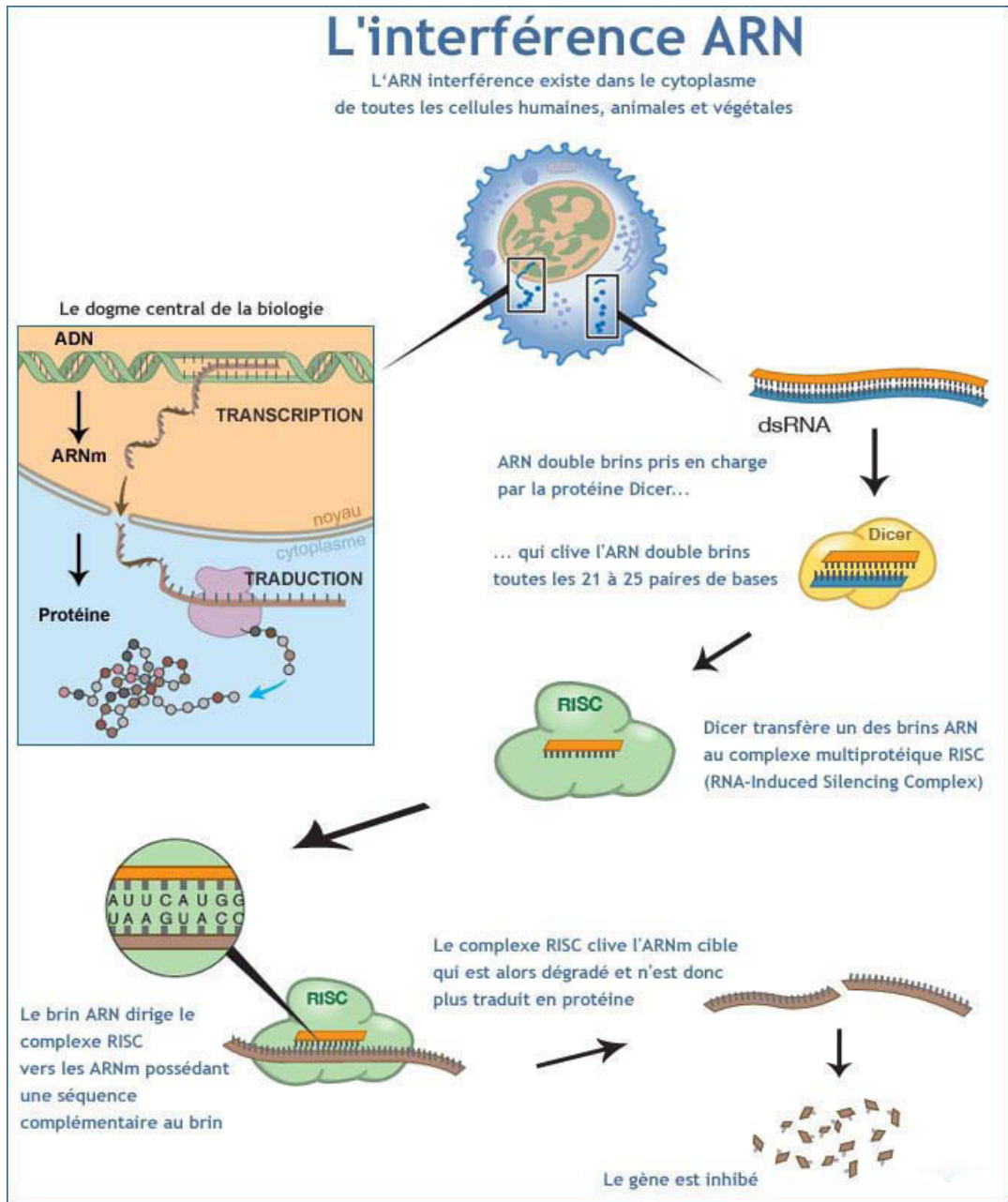


Figure 53 : Principe de l'ARN interférence. Seul le processus d'extinction post-transcriptionnel est représenté dans cette figure. Le mécanisme d'entretien de l'extinction au niveau transcriptionnel n'est pas représenté ici.

coloration des fleurs de petunia. Afin d'intensifier la couleur pourpre des pétales, Jorgensen et ses collaborateurs avaient introduit une copie du gène de la chalcone synthase dans des pétunias. De manière intéressante, certains pétunias devenaient partiellement ou totalement blancs ; l'introduction du transgène chalcone synthase ayant inhibée l'expression du gène endogène. Ce mécanisme fut alors observé de nouveau lors d'études d'introduction d'ARN double-brin chez le nématode *Caenorhabditis elegans*, par Fire et ses collaborateurs (1998) et appelé ARN interférence.

Au cours de l'évolution, ce mécanisme a tout d'abord probablement été sélectionné pour prémunir les cellules eucaryotes primitives d'attaques virales (Baulcombe, 2004). Dans les plantes, ce processus d'interférence ARN s'est ensuite diversifié en divers mécanismes qui permettent, outre son rôle dans la défense virale, de protéger le génome des transposons et de réguler l'expression de gènes. La validation fonctionnelle de gènes par ARN interférence repose sur l'un de ces mécanismes naturels: l'extinction de gènes par des petits ARN interférants (siRNA) dans le cytoplasme (Figure 53). Afin d'éteindre spécifiquement l'expression d'un gène, une construction est créée à partir de la séquence partielle du transcrit mature du gène ciblé, afin de former des produits ARN double-brin en structure en épingle à cheveux, homologue à la séquence du gène ciblé. Cette construction est introduite dans l'organisme et les ARN double-brins formés sont détectés par la cellule comme étant des produits aberrants. Ils sont alors clivés par des enzymes, appelées Dicer-like chez les plantes, en courts et longs siRNA double-brin, d'approximativement 21 et 24-mer respectivement. Ces 2 classes de siRNA sont supposées avoir des fonctions distinctes : les siRNA de 21-mer conduisent directement le signal de dégradation post-transcriptionnelle des ARN messagers, et les siRNA de 24-mer entretiennent le signal d'extinction post-transcriptionnelle et permettent une extinction transcriptionnelle de l'expression du gène par méthylation de l'ADN au locus du gène homologue. L'un des 2 brins des courts siRNA va être pris en charge par un complexe protéique appelé RISC (RNA-induced silencing complex), et se lier à l'ARN messager du gène d'intérêt par complémentarité, qui va être reconnu et dégradé. Le signal d'extinction du gène est alors entretenu par les longs siRNA : ils amplifient le signal en se liant aux ARNm endogènes, qui sont alors convertis en double-brin par l'enzyme polymérase RdRP (RNA-directed RNA polymerase), puis reconnus et clivés pour former une population plus importante de siRNA. Une méthylation de l'ADN au locus du gène ciblé peut de plus avoir lieu afin de maintenir une inactivation de l'expression du gène. L'extinction du gène persiste à travers les divisions cellulaires et peut être transmise aux générations suivantes.

L'ARN interférence est largement utilisée maintenant pour la validation fonctionnelle de gènes et est notamment très utile pour étudier la fonction de gènes chez des variétés polyploïdes, comme le blé tendre, pour lesquelles l'obtention de mutants perte-de-fonction est difficile et longue (Yan *et al*, 2004 ; Loukoianov *et al*, 2005 ; Regina *et al*, 2006 ; Travella *et al*, 2006 ; Uauy *et al*, 2006). De plus, l'ARN interférence peut permettre de générer un panel de plantes transgéniques présentant un taux d'extinction variable, très utile lorsque la complète extinction d'un gène est létale par exemple (Travella *et al*, 2006 ; Fu *et al*, 2007). Fu et ses collaborateurs (2007) ont recensé dans une revue différents points importants pour la validation fonctionnelle par ARN interférence chez le blé tendre

- *L'identité de la séquence de la cassette RNAi*: Pour que la dégradation post-transcriptionnelle spécifique d'ARN messagers soit déclenchée, il faut qu'il y ait des « bandes » d'identité parfaite en nucléotides entre la séquence insérée dans la cassette et la (les) séquence(s) de la population de transcrit(s) mature(s) ciblé(s). Un minimum de 21 nt d'identité de séquence continue parfaite est généralement requis. Cependant, des études ont montré que dans certains cas, ce minimum n'est pas toujours suffisant pour produire efficacement une extinction du gène ciblé (Miki *et al*, 2005 ; Yue *et al*, 2007 ; McGinnis *et al*, 2007).
- *La région de la séquence d'intérêt sélectionnée pour la cassette RNAi* : Il est important de sélectionner une région du transcrit ciblé exclusive à ce gène et ses homéologues, afin de ne pas éteindre par exemple d'autres gènes appartenant à une même famille de gènes. Cependant, si l'objectif est d'éteindre l'ensemble des membres d'une famille de gène en particulier, il peut être intéressant de choisir une région conservée entre ces membres exclusive à cette famille de gènes.
- *La taille de la séquence sélectionnée pour la cassette RNAi* : Une taille minimale de la séquence insérée est requise pour une extinction efficace de l'expression d'un gène. Beaucoup d'études de validation fonctionnelle par RNAi chez le blé tendre reportent l'utilisation de séquences d'intérêt insérée dans la cassette ayant une taille allant de 200 à 550 pb. Bien que chez *Nicotiana benthamiana* l'utilisation d'ARN double-brin de 23 nt a permis d'induire efficacement la dégradation d'ARN messagers ciblés, il semblerait qu'un minimum de 120 pb soit préférable chez le blé tendre. La taille maximale requise est quant à elle plus floue. Des études ont reporté des succès

d'extinction de gènes par RNAi avec une séquence de 683 pb chez le blé tendre (Li *et al*, 2005) et jusqu'à 1215 pb chez l'orge (Holzberg *et al*, 2002).

- *L'efficacité de l'extinction et le nombre d'évènements transgéniques* : Le taux d'extinction obtenu est très variable entre les études et semble gène-dépendant. Des taux d'extinction allant de 29 à 100% ont été reportés. Le nombre d'évènements transgéniques indépendants nécessaires pour conclure sur la fonction d'un gène n'est pas clairement établi, mais il est raisonnable de penser que le phénotype mutant doit être retrouvé dans plusieurs évènements transgéniques indépendants.

De plus, quelques autres points importants ont été observés dans plusieurs études :

- Les plantes transgéniques homozygotes pour le transgène ont tendance à présenter des taux de réduction d'accumulation du transcrit ciblé plus importants et des phénotypes plus marqués que des plantes transgéniques hétérozygotes, du fait d'une accumulation plus importante de siRNA, suggérant un effet dose de l'ARN interférence (Travella *et al*, 2006 ; Fu *et al*, 2007).
- L'introduction de séquences introniques ou « espaceurs » entre les séquences d'intérêt insérées en orientation sens et antisens dans la cassette RNAi permet d'accroître considérablement l'efficacité de l'extinction de gènes, aidant probablement à la formation de la structure ARN double-brin en épingle à cheveu (Smith *et al*, 2000).
- Une variation dans la taille de l'intron et de la séquence d'intérêt dans la cassette RNAi peut influencer l'extinction du gène et donc le phénotype mutant (Smith *et al*, 2000 ; Hirai *et al*, 2007).

La construction de la cassette RNAi-*TaGW2* et l'obtention des lignées transgéniques T0 par transformation d'embryons immatures de blé tendre par biolistique ont été réalisées par la plate-forme de validation fonctionnelle de l'UMR.

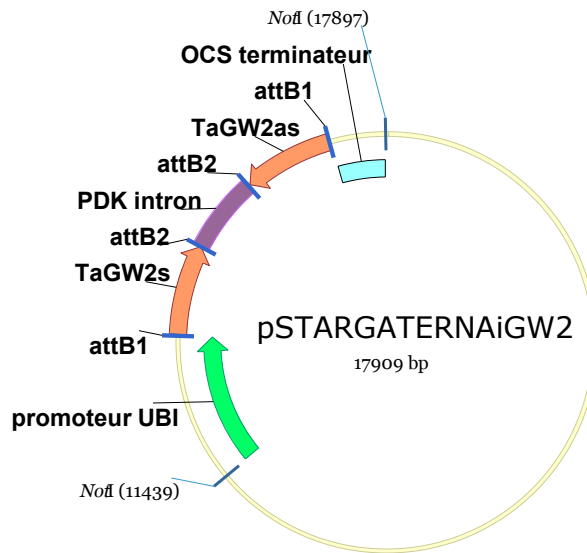


Figure 54 : Représentation schématique de la cassette RNAi *TaGW2* dans le vecteur pSTARGATE modifié. TaGW2s et as: séquence partielle du CDS de *TaGW2* (des bases 34-1275) en orientation sens et antisens ; PDK intron : intron du gène de la kinase pyruvate déshydrogénase d'*Arabidopsis* ; promoteur UBI : promoteur de l'ubiquitine. Les positions des sites NotI sur le vecteur recombinant sont indiquées.

10.1 Construction de la cassette RNAi-*TaGW2*

La construction RNAi-*TaGW2* a été faite à partir d'une adaptation du vecteur pSTARGATE, fourni par le CSIRO (Australie ; <http://www.pi.csiro.au/rnai/vectors.htm>; Figure 54). En effet, le vecteur d'origine a été modifié pour permettre la construction d'une cassette RNAi et d'une cassette de sélection indépendantes. Un fragment de 1242 pb de la séquence codante de *TaGW2-A* (de la base 34 à 1275 du CDS de *TaGW2-A*) a été amplifié par PCR sur le vecteur recombinant pGEMT-TaGW2-A à l'aide du couple d'amorces AttbTaGW2-FL-CDS, purifié puis cloné dans le vecteur donneur pDONR221 par une réaction BP (technologie Gateway). La séquence partielle de *TaGW2-A* a ensuite été sous-clonée en orientation sens et antisens dans le vecteur pSTARGATE modifié, par des réactions LR (technologie Gateway). Le vecteur recombinant obtenu a été digéré par l'enzyme de restriction NotI et la cassette RNAi-TaGW2 a été purifiée par migration sur gel suivie de lavages au butanol et à l'éther (protocole confidentiel – propriété de l'INRA).

La cassette de sélection contient le gène de la phospho-mannose isomérase sous le contrôle du promoteur constitutif fort de l'ubiquitine de maïs, pour une sélection des plantes transformées sur un milieu contenant uniquement du mannose comme source carbonée.

10.2 Transformation d'embryons immatures de blé tendre par biolistique et régénération de plantules

Cette partie a été réalisée par le personnel de la plateforme « transgène blé » de l'UMR. La cassette RNAi-TaGW2 et la cassette de sélection ont été mélangées à un ratio molaire de 2 :1, puis mélangées à des billes d'or de 0.6µm de diamètre (BIORAD). Les procédures de transformation et de régénération ont été effectuées selon le protocole décrit par Wright *et al* (2001) et Pellegrineschi *et al.* (2002), avec quelques modifications. Brièvement, ce mélange de cassettes est bombardé sur des embryons immatures de blé tendre à l'aide d'un canon à particule (BioRad PDS-1000/He biolistic particle system, Du Pont). Les embryons sont alors laissés 24h sur un milieu plasmolysant (annexe 1), puis transférés pendant 15 jours sur un milieu de callogénèse sans sélection (annexe 1), à l'obscurité. Les cals sont alors repiqués successivement sur 3 différents milieux de

régénération sélectifs (annexe 1), en présence de lumière tamisée puis vive. Les plantules vigoureuses développant des départs de feuilles sont repiquées sur un milieu d'enracinement (annexe 1), et placées de nouveau à la lumière. Lorsque les plantules présentent des systèmes racinaires et foliaires suffisamment développés, elles sont repiquées en terre dans des godets puis en pots de 12L et placées en chambre transgénique. Des sachets d'auto-fécondation sont placés sur les épis fertiles pour la génération de graines T1.

10.3 Génotypage des plantes transformées.

Les plantes transformées ont été génotypées pour la présence de la cassette par PCR sur ADN génomique, à l'aide du couple d'amorces PDK-intron qui permet une amplification de 387 pb de l'intron du gène PDK du transgène. L'amplification a été réalisée sur 100 ng d'ADN génomique, selon le programme suivant : 95°C, 3 min puis 40 cycles de 95°C, 30 sec ; 55°C, 30 sec ; 72°C, 30 sec.

10.4 Quantification du taux d'expression endogène de *TaGW2* dans les plantes transformées.

Les taux d'expression endogène du gène *TaGW2*, toutes copies confondues, et des copies homéologues *TaGW2-A*, *TaGW2-B* et *TaGW2-D* ont été quantifiés par PCR quantitative en temps réel dans des feuilles des plantes transgéniques et null-ségrégantes, à l'aide des couples d'amorces RNAiTaGW2endo, RNAiTaGW2Aendo, RNAiTaGW2Bendo et RNAiTaGW2Dendo, respectivement, comme mentionné en section 3.2.3. L'amorce sens du couple RNAiTaGW2endo est située dans les premiers nucléotides de la séquence de *TaGW2*, partie manquante dans le transgène et permet donc une amplification exclusive des transcrits *TaGW2* endogènes, et non du transgène.

10.5 Phénotypage des plantes transformées.

Pour les 2 évènements de transformation, le phénotypage des plantes transgéniques et nulle-ségrégantes a été réalisé à la génération F2. Nous avons choisi de ne pas le mener à la génération F1, du fait que certaines plantes transgéniques et témoins présentaient un

Transgénique



Contrôle



Figure 55 : Phénotypes retrouvés sur des plantes contrôle et transgéniques RNAi en génération F1. Il y a formation d'épillets surnuméraires.

phénotype étrange de formation de fleurs surnuméraires sur des épillets (Figure 55), résultant sans doute du stress de la transformation subi par les plantes.

La croissance et le phénotypage des plantes F2 de chaque évènement de transformation ont été conduits séparément dans le temps. La conduite des plantes F2 de l'évènement T105 a eu lieu de juin 2010 à janvier 2011, et celle de l'évènement T111 de novembre 2010 à Mai 2011.

Les épis de toutes les talles de chaque plante ont été étiquetés à anthèse ; la date a été enregistrée lorsque les premiers épillets de chaque épi fleurissent. Pour les 2 évènements transgéniques, les épis ont été prélevés à maturité physiologique (1000°CJours après anthèse). Le nombre d'épillets par épi et le nombre de grains par épi ont alors été déterminés. Les mesures individuelles de grains suivantes ont été réalisées sur les 2 grains basaux des 2 épillets centraux de l'épi du maître-brin de chaque plante : les masses individuelles fraîche et sèche du grain (MF et MS, respectivement) ; la masse en eau du grain à maturité (ME); le volume (V), la longueur (Lo), la largeur (La) et l'épaisseur (E) du grain. Les valeurs de MF, V, Lo, La et E ont été obtenues à partir de grains fraîchement récoltés et les valeurs de MS après déshydratation des grains dans une étuve à 80°C, pendant 48h. La masse en eau du grain à maturité est la différence entre MF et MS du grain. Le volume du grain a été estimé par une méthode de déplacement de l'eau à l'aide d'un pycnomètre. Les dimensions du grain (longueur, largeur, épaisseur) ont été déterminées à l'aide d'un système photographique assisté par ordinateur. Tous les grains restants de l'épi ont été comptés, regroupés, pesés et broyés en farine complète à l'aide d'un moulin (6800 Cyclotec mill). Le contenu total en azote (en % de la MS) a été déterminé sur chaque mouture suivant la méthode de combustion de Dumas modifiée (méthode AOAC n°992.23) à l'aide d'un analyseur FlashEA 1112N/Protein analyser (Thermo Scientific). Chaque extrait a été analysé en duplicat. Le contenu total en protéines a été calculé en multipliant le contenu total en azote par 5,7 (Sosulski et Imafidon, 1990). Les contenus en azote et en protéines par grain sont les masses totales d'azote et de protéines déterminées dans la masse totale de la mouture complète, divisées par le poids moyen d'un grain de l'épi utilisé pour la mouture.

Dans la lignée T105, les grains basaux d'épillets centraux d'épis de talles secondaires ont été récoltés à 260°CJours. Ce temps thermique est considéré comme le stade de développement du grain auquel la division cellulaire s'arrête dans la variété Récital, sous des conditions environnementales similaires aux conditions de croissance appliquées aux plantes transgéniques au cours de ces expériences (Nadaud *et al*, 2010). Le nombre de

cellules de l'albumen a été compté selon la méthode décrite par Rijven et Wardlaw (1966), puis modifiée par Singh et Jenner (1982). Brièvement, les noyaux des cellules de l'albumen sont colorés par une réaction de Feulgen et 10 aliquots de 10 μ l de la suspension de cellules de l'albumen préparée sont déposés sur une membrane de nitrocellulose (0,22 μ m de diamètre, Millipore). Le nombre de noyaux est alors compté automatiquement à l'aide d'une cellule de comptage macro assistée par ordinateur (logiciel ImageJ). Ce comptage de cellules de l'albumen à 260°CJours n'a pu être réalisé sur les plantes F2 de la lignée T111 du fait du très faible nombre de talles produites par l'ensemble des plantes de cette lignée.

11 Sur-expression in planta de *TaGW2-A*

La construction de la cassette de sur-expression et l'obtention des lignées transgéniques T0 par transformation d'embryons immatures de blé tendre par biolistique ont été réalisées par la plate-forme de validation fonctionnelle de l'UMR.

11.1 Construction de la cassette de surexpression Bx7-*TaGW2A*-PMI

La cassette de surexpression a été construite dans un vecteur à entrées multiples par une réaction LR (technologie Gateway), à partir de 3 plasmides recombinants d'entrée :

- un plasmide contenant le promoteur fort graine-spécifique Bx7 d'une gluténine de haut poids moléculaire.
- un plasmide contenant l'ADNc complet de *TaGW2-A*
- un plasmide contenant un terminateur du gène de la nopaline synthase, suivi du gène de sélection PMI sous le contrôle du promoteur fort de l'ubiquitine, puis d'un terminateur du gène de la nopaline synthase.

Contrairement à la cassette RNAi, la cassette de sur-expression possède physiquement le gène de sélection PMI, qui est donc intégré dans les plantes transformées en même temps que la cassette Bx7-*TaGW2A*-PMI.

11.2 Obtention des lignées de sur-expression *TaGW2-A*

Le vecteur recombinant final a été digéré par l'enzyme de restriction NotI et la cassette Bx7-*TaGW2A*-PMI a été purifiée par migration sur gel suivie de lavages au butanol et à

l'éther (protocole confidentiel – propriété de l'INRA). Des embryons immatures de *T.aestivum* cv Bobwhite ont été transformés par biolistique avec la cassette Bx7-*TaGW2A*-PMI et les plantules T0 ont été régénérées, comme précédemment mentionné en section x.

12 La technique de double-hybride

Le double-hybride est une méthode de détection d'interactions protéiques *in vivo* dans la levure. A l'heure actuelle, c'est l'une des techniques les plus employées, avec le pull-down, pour la détection d'interactions protéine-protéine (IPP). Le double-hybride présente de nombreux avantages mais également des inconvénients à prendre en considération au cours de ce type d'analyse (Brückner *et al*, 2009). Les avantages majeurs du double-hybride sont son accessibilité à moindre coût (pas d'équipements de pointe requis, seulement du matériel de techniques microbiologiques) et la rapidité de détection d'IPP. De plus, cette technique est adaptée à des cribles haut-débit d'IPP. Malheureusement, le double-hybride souffre également des inconvénients suivants : il ne permet que la détection d'interactions binaires et pas la détection d'interaction simultanée de multiples protéines, comme dans le cas d'un complexe par exemple. De plus, ce type d'analyse génère un taux élevé de faux-positifs, et requiert des confirmations des interactions potentielles révélées par d'autres techniques biochimiques de détection d'IPP.

Historiquement, les interactions en double-hybride étaient d'abord détectées dans le noyau de levures. La limitation des classes de protéines pouvant être testées pour des IPP dans ce type de double-hybride (appelé classique ou nucléaire) a poussé à l'émergence ces dernières années de nouveaux systèmes de double-hybride (Brückner *et al*, 2009). Dans ce travail de thèse, la recherche d'IPP pour *TaGW2-A* a été entreprise sans *a priori* par criblage de banque ADNc par double-hybride nucléaire (système Matchmaker, Clontech) et cytoplasmique (système DUALHunter, Dualsystems Biotech). La séquence maximale de *TaGW2-A* n'autoactivant pas les gènes rapporteurs dans un système nucléaire a préalablement été recherchée grâce à un protocole dérivé du système de double-hybride nucléaire, le système ProQuest (Invitrogen).

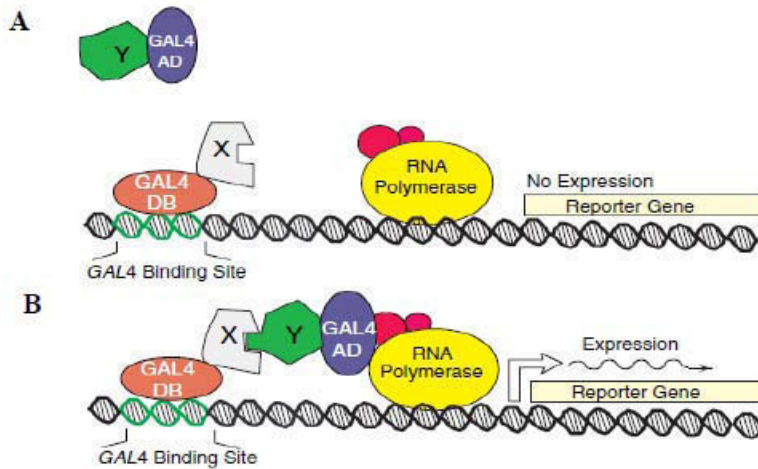


Figure 56 : Principe du double-hybride nucléaire. Exemple de la modularité du facteur de transcription GAL4. Deux protéines chimères sont créées : la protéine X (appât) est fusionnée au domaine de liaison à l'ADN (DBD) de GAL4 ; la protéine Y (proie) au domaine d'activation (AD) de GAL4. La souche de levure employée dans ce système double hybride nucléaire possède des gènes rapporteurs sous le contrôle d'un promoteur possédant une région régulatrice (UAS) fixatrice du domaine DBD de GAL4. **(A)** Si les protéines X et Y n'interagissent pas, il n'y a pas de rapprochement du domaine AD de GAL4 au promoteur du gène rapporteur, et le gène rapporteur n'est donc pas exprimé dans la levure. **(B)** Si les protéines X et Y interagissent, le domaine AD de GAL4 peut alors recruter la machinerie basale de transcription ; le gène rapporteur est exprimé dans la levure.

12.1 Le double-hybride nucléaire

12.1.1 Principe général du double-hybride nucléaire en levure

Le double-hybride nucléaire repose sur la modularité du facteur de transcription GAL4 ou LexA (Figure 56). Les facteurs de transcription sont constitués d'un domaine de liaison à l'ADN (DBD), et d'un domaine d'activation de la transcription (AD) capable de recruter la machinerie basale de transcription. Une protéine appât, dont on veut identifier les interacteurs, est fusionnée au DBD du facteur de transcription. L'interacteur potentiel, appelé protéine proie, est fusionné au AD du facteur de transcription. Les protéines fusions sont co-exprimées dans le noyau des cellules de levure contenant des gènes rapporteurs dont l'expression est placée sous le contrôle du site de fixation pour le DBD du facteur de transcription. Lorsque la protéine proie est capable d'interagir avec la protéine appât, il y a rapprochement dans l'espace des deux domaines du facteur de transcription et donc reconstitution d'un facteur de transcription fonctionnel. Le DBD se fixe sur son site de fixation dans le promoteur du gène rapporteur, et le domaine d'activation du facteur de transcription recrute et amène la machinerie basale de transcription sur le promoteur du gène rapporteur. Il y a alors transcription du gène rapporteur.

12.1.2 Le système ProQuest (Invitrogen)

Des fragments et la séquence complète codante de TaGW2-A ont été clonés dans le vecteur pDEST32, en C-terminale et dans le cadre de lecture du DBD de GAL4, comme mentionné en p. 68. Des levures de souche MaV203 ont été rendues compétentes puis ont été co-transformées avec le vecteur pDEST22 vide, contenant le AD de GAL4, et l'une des constructions protéine fusion DBD GAL4-TaGW2, à l'aide du kit S.c. EasyComp Transformation kit (Invitrogen), selon les recommandations du fournisseur. Par protéine fusion appât considérée, plusieurs clones de levure co-transformés ont été testés pour l'autoactivation de la protéine appât sur des milieux de sélection pour l'expression des divers gènes rapporteurs, selon les recommandations fournies dans le protocole du système ProQuest d'Invitrogen. Une protéine fusion DBD GAL4-appât est dite autoactivatrice dans ce système lorsque la colonie de levure co-transformée testée est bleue après un test X-Gal sur membrane de nylon (sélection pour l'expression du gène rapporteur LacZ), et pousse sur les milieux de sélection pour les gènes rapporteurs HIS et URA (milieux $-L/-T/-H + 3AT$ et $-L/-T/-U$, respectivement ; Annexe 1).

Dans le système ProQuest, trois constructions contrôles sont fournies :

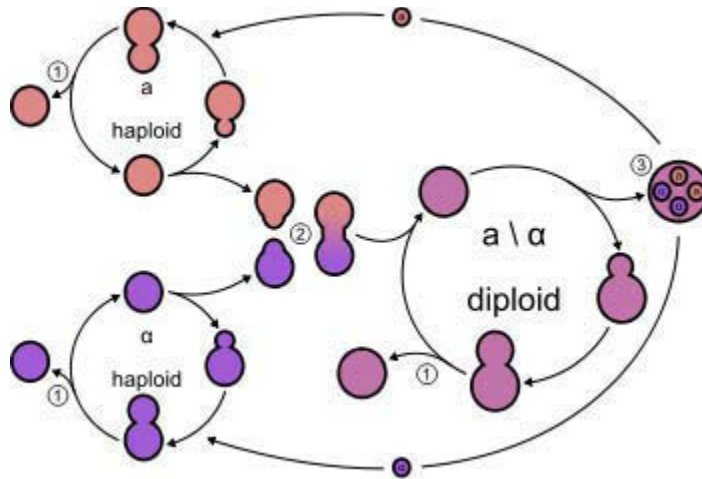


Figure 57 : Cycle de vie de *S. cerevisiae*. La levure est capable de se multiplier sous deux formes : une forme diploïde ($2n = 32$) et une forme haploïde ($n = 16$). Les cellules haploïdes se multiplient en bourgeonnant (1) : la cellule mère bourgeonne une cellule fille plus petite par mitose, possédant la même information génétique. Il existe des cellules haploïde de type sexuel « a » et « α ». ces deux types de cellules ne se distinguent pas morphologiquement mais par la phéromone qu'elles produisent : MATa ou MATα. Les phéromones libérées permettent l'amorce du processus de fécondation (2). La fusion entre une levure haploïde de type « a » et une levure haploïde de type « α » donne naissance à une cellule diploïde « a/α ». Tant que l'environnement est favorable, le diploïde se multiplie par bourgeonnement (1, dans le cycle diploïde). Si les nutriments viennent à manquer, la levure diploïde repasse en phase haploïde par un processus de méiose (3).

Dans le système de criblage de banque ADNc par Matchmaker (Clontech), l'étape appelée « mating » correspond à l'étape (2) de ce schéma.

- la protéine Krev1 fusionnée au DBD GAL4 (p32-Krev1) ;
- la version sauvage de la protéine RalGDS, interacteur fort de Krev1, fusionnée au AD GAL4 (p22-RalGDSwt) ;
- et une version mutante (L829P) de la protéine RalGDS incapable d'interagir avec Krev1, fusionnée au AD GAL4 (p22-RalGDSm2).

Les couples p32-Krev1/ p22-RalGDSwt et p32-Krev1/ p22-RalGDSm2 constituent donc des contrôles positifs et négatifs respectivement des tests de l'expression des gènes rapporteurs, dans le système ProQuest (Invitrogen).

12.1.3 Le système Matchmaker (Clontech)

Dans le système Matchmaker, les banques ADNc sont construites dans le vecteur pGADT7-Rec, à l'aide du kit Make your own « Mate & PlateTM » Library System (Clontech) selon les recommandations du fournisseur ; et hébergées dans la souche de levure Y187. La mise en compétence et la transformation de la souche de levure Y187 sont réalisées à l'aide des kits S.c. EasyCompTM Transformation (Invitrogen) et YeastmakerTM Yeast Transformation System 2 (Clontech), respectivement, selon les recommandations des fournisseurs. La protéine fusion DBD GAL4-appât est construite dans le vecteur pGBKT7 ; et la souche de levure Y2HGold est rendue compétente à l'aide du kit S.c. EasyCompTM Transformation (Invitrogen) puis transformée avec la construction DBD GAL4-appât à l'aide du kit YeastmakerTM Yeast Transformation System 2 (Clontech) selon les recommandations du fournisseur. Le criblage de la banque ADNc avec la protéine fusion appât est réalisé par « mating » (rencontre sexuelle ; Figure 57) des deux souches de levures selon le protocole décrit dans le manuel de l'utilisateur du MatchmakerTM Gold Yeast Two-hybrid System, avec quelques modifications. Les levures diploïdes formées présentant des interactions positives entre la protéine appât et une proie potentielle sont criblés sur des milieux sélectifs pour les gènes rapporteurs contenus dans la souche de levure Y2HGold, selon les recommandations du fournisseur. Pour rappel, la souche Y2HGold possède les gènes rapporteurs AUR1-C, HIS3, ADE2 et MEL1.

Construction des banques ADNc dans la souche Y187

Les deux banques d'ADNc générées dans le cadre de ce travail de thèse, nommées 'Division' et 'Remplissage', ont été construites à partir d'ARN totaux de grains de blé entiers

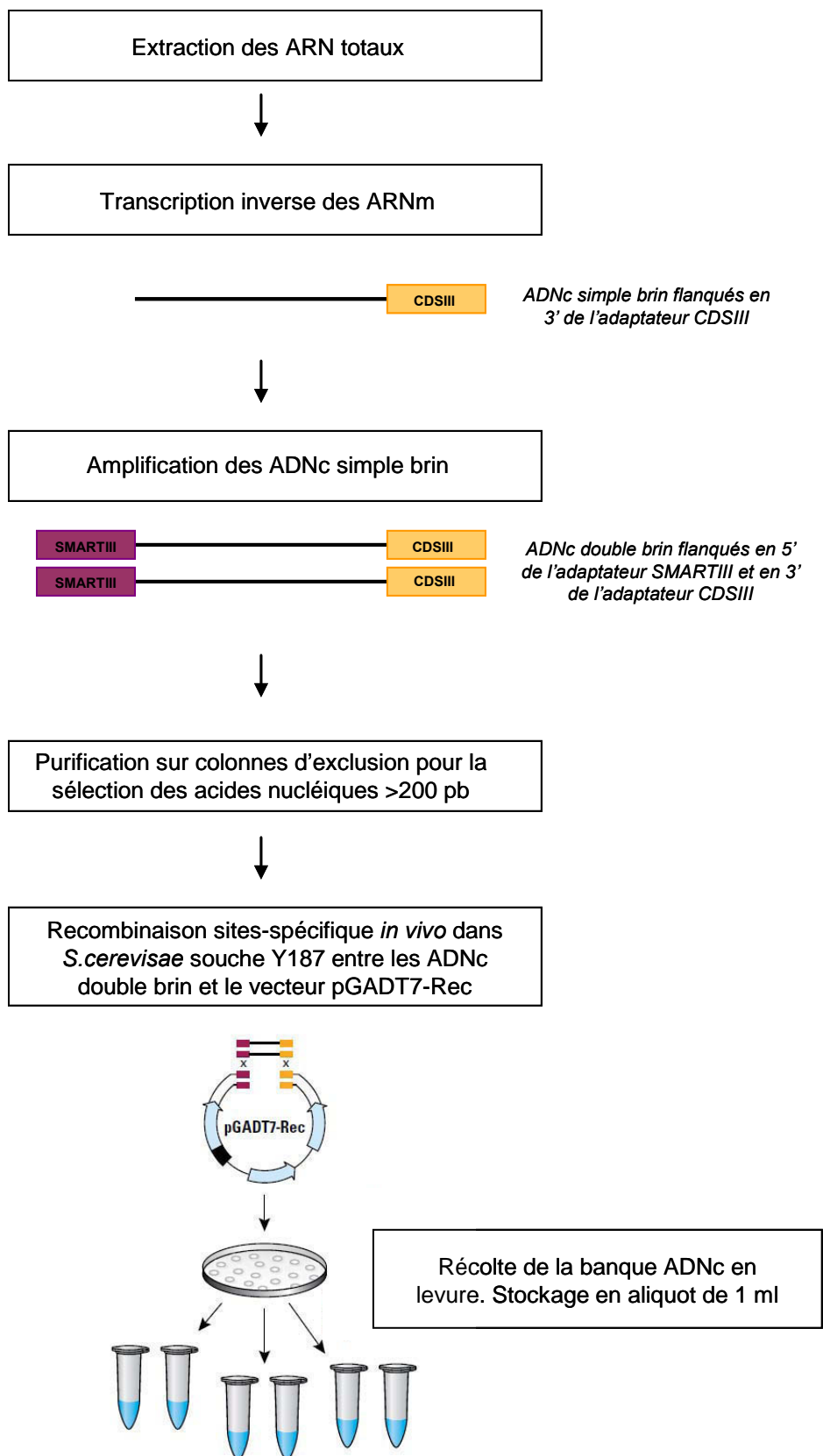


Figure 58 : Schéma simplifié du protocole expérimental de création d' une banque ADNC en levures Y187, dans le système Matchmaker. D' après le manuel de l' utilisateur du kit Make your own « Mate & Plate » Library system (Clontech).

à l'aide du kit « Make your own Mate & Plate Library System » (Clontech). Un schéma du protocole expérimental d'une construction d'une banque ADNc à l'aide de ce kit est donné en Figure 58. Seuls quelques détails propres à nos expérimentations et les modifications apportées à ce protocole seront mentionnés ci-après.

Pour la construction de la banque nommée 'Division', des grains récoltés à 80, 160 et 220°J ont été rassemblés et broyés ensemble à l'azote liquide dans un mortier préalablement refroidi à l'azote liquide. Pour la construction de la banque nommée 'Remplissage', des grains récoltés à 300, 500 et 600°J ont été utilisés et broyés comme décrit précédemment. Les ARN totaux ont été extraits par une extraction ARN au tampon Borate comme décrit p.X.

Brièvement, la synthèse d'ADNc simple brin a été réalisée à l'aide de 2µg d'ARN totaux et de 10pmoles d'amorce oligodT 'CDS III', selon les recommandations du fournisseur. L'amplification des ADNc par PCR Longue-Distance a ensuite été entreprise sur 25 cycles ; et la purification des ADNc double-brin a été réalisée sur des colonnes CHROMA SPINTM TE-400, afin de ne sélectionner que les molécules de taille supérieure à 200 pb.

Les cellules de levures Y187 ont été rendues compétentes à l'aide des solutions I et II du kit « S.c. EasyCompTM Transformation kit » commercialisé par Invitrogen. Une colonie Y187 sur milieu YPAD solide est repiquée dans 10 ml de milieu YPAD liquide. La solution est vortexée pendant 30s afin de bien disperser les cellules de levures et incubée pendant 12h à 30°C sous agitation 225 rpm. Le titrage de la culture de levure Y187 est alors effectué comme suit. La suspension est vortexée pendant 30s et 20µl de suspension sont mélangés à 2 ml d'eau dans une cuve à spectrophotométrie. Le titrage est déterminé par une lecture de densité optique à 600nm, sachant qu'une $DO_{600nm} = 0.1$ correspond à 1.10^6 cellules/ml. Dans 10 ml de milieu YPAD liquide sont alors ensemencés $0,5.10^8$ cellules et placés à 30°C sous agitation 225 rpm jusqu'à obtenir un titrage de la culture de 2.10^7 cellules/ml. La culture est alors centrifugée 5 min à 500g à température ambiante. Le surnageant est éliminé et le culot est repris dans 10 ml de solution I. Suit une nouvelle centrifugation de 5 min, à 500g, à température ambiante. Le surnageant est éliminé et le culot est finalement repris dans 600µl de solution II. Les cellules Y187 maintenant compétentes peuvent être utilisées immédiatement ou stockées à -80°C enveloppées dans plusieurs épaisseurs de papier absorbant.

La construction de banque ADNc dans les cellules de levures Y187 repose sur un clonage par recombinaison *in vivo* directement dans la levure entre les séquences adaptatrices flanquant les ADNc double brin (SMART III et CDS III) et des séquences homologues dans le plasmide pGADT7-Rec (Clontech). Pour rappel, le plasmide pGADT7-Rec permet une fusion du domaine AD GAL4 en N-terminal des ADNc produits à partir d'une population d'ARN d'intérêt. Le mode opératoire du fournisseur pour la transformation de la souche de levure Y187 à grande échelle (« library-scale ») a été modifié pour une meilleure efficacité de transformation. Vingt microlitres d'ADNc double brin préalablement purifiés, 200 µg d'ADN cargo simple brin et 3µg de plasmide pGADT7-Rec sont ajoutés à 600 µl de cellules de levures Y187 compétentes, sur la glace. Le mélange est vortexé vigoureusement pendant 30 s puis additionné de 2,5 ml d'une solution de PEG/LiAc (Annexe 1). Le mélange est de nouveau vigoureusement vortexé puis incubé à 30°C pendant 45 min sous agitation 200 rpm, avec un léger vortexage des cellules toutes les 15 min. Soixante-dix microlitres de DMSO sont ajoutés et le tube est inversé plusieurs fois doucement. Il est important de ne pas vortexer à cette étape. Un choc thermique des cellules est alors effectué à 42°C pendant 20 min avec un léger vortexage des cellules toutes les 10 min puis le tube est refroidi sur la glace pendant 2 min. Les cellules de levures Y187 transformées sont centrifugées à 700g, pendant 5 min. Le surnageant est éliminé et le culot est repris dans 3 ml de milieu YPD. Les cellules sont ensuite incubées 90 min à 30°C sous agitation 200 rpm. Les cellules Y187 transformées sont centrifugées 5 min à 700g et le culot est finalement repris dans 15 ml d'une solution NaCl à 0,9% (p/v). Cent microlitres de cellules transformées diluées au 1/10^{ème} et 1/100^{ème} sont étalées sur milieu SD/-L et incubées pendant 3 jours à 30°C.

Le titrage de la banque réalisée est déterminé comme suit :

Nombre de clones indépendants dans la banque = Nbre de cfu/ml sur une boîte SD/-L x volume de resuspension finale (en ml).

La banque doit contenir au minimum 1 million de clones indépendants pour pouvoir poursuivre les étapes suivantes.

Le restant des cellules Y187 transformées est étalé sur des boîtes SD/-L de 15 cm de diamètre (environ 260 boîtes), à raison de 50 µl par boîte, et incubé à 30°C pendant 4 jours. Les boîtes sont ensuite placées à 4°C pendant 4h. Deux fois un ml d'une solution de stockage (YPAD /

25% Glycérol) sont ajoutés successivement sur chacune des boîtes. Les cellules de levures Y187 transformées sont alors détachées de la boîte à l'aide de billes de verre stériles. Les mêmes billes sont utilisées pour détacher l'ensemble de la banque. Les suspensions de levures de toutes les boîtes sont alors récoltées (soit environ 400 ml) et assemblées dans un même flacon stérile. La densité cellulaire de la banque de levures est alors déterminée avec une cellule de comptage de Fuchs-Rosenthal. La densité cellulaire doit être au minimum de 2.10^7 cellules /ml. Si elle est inférieure à cette valeur, le volume de la suspension est réduit par une centrifugation. La banque d'ADNc en cellules de levures Y187 est alors aliquotée par 1 ml pour une conservation à court-terme et par 50 ml pour une conservation à long-terme, et stockée à -80°C .

Construction de la protéine fusion appât dans la souche Y2HGold

L'ADNc du gène d'intérêt ou un fragment, est cloné dans le vecteur pGBKT7 à l'aide du kit In-Fusion® Advantage PCR Cloning kit (Clontech) selon les recommandations du fournisseur, puis entreteu dans *E.coli*. Brièvement, le vecteur pGBKT7 est tout d'abord linéarisé par une double digestion enzymatique *Bam*HI et *Eco*RI pendant une nuit à 37°C . L'intégralité du mélange réactionnel est déposé sur gel puis purifié avec le kit Nucleospin® Extract II Kit (Macherey-Nagel), selon le protocole du fournisseur, avec une élution finale dans du tampon d'élution préchauffé à 70°C .

L'ADNc d'intérêt est amplifié par PCR à l'aide d'amorces contenant 24 pb d'homologie de séquence avec les extrêmités de l'ADNc d'intérêt et 16 pb d'homologie de séquence avec les extrêmités linéaires du vecteur pGBKT7. L'intégralité du produit d'amplification est déposé sur gel et purifié avec le kit Nucleospin® Extract II Kit (Macherey-Nagel), selon les recommandations du fournisseur. Les produits PCR purifiés sont dosés sur gel d'agarose 2%, à l'aide d'une échelle de poids moléculaire de quantification.

Pour la réaction de ligation, sont mélangés dans un tube 100ng de vecteur pGBKT7 linéarisé, 50ng de produit PCR purifié, 2µl de 5X In-Fusion Reaction Buffer, 1µl d'enzyme In-Fusion, et de l'eau déionisée pour un volume final de 10µl. Il est important de respecter un ratio molaire de 2 :1 pour le mélange du vecteur pGBKT7 linéarisé et du produit d'amplification. Le mélange est incubé 15 min à 37°C puis 15 min à 50°C et finalement refroidi sur la glace pendant 2 min. Le volume réactionnel est porté à 50µl avec du tampon TE (pH 8).

Cinquante microlitres de bactéries chimiquement compétentes (Invitrogen) sont transformés avec 2,5 µl de réaction de ligation In-Fusion, selon les recommandations du

fournisseur. La construction est vérifiée par séquençage avec l'amorce universelle T7 sens (Beckman Coulter Genomics), et le plasmide recombinant est finalement extrait à l'aide du kit PureLink QuickPlasmid (Invitrogen).

La souche de levure Y2HGold est rendue compétente comme décrit précédemment pour la souche Y187. Le culot final est repris dans 1 ml de solution II et les cellules Y2HGold sont aliquotées par 50µl.

Les cellules Y2HGold sont transformées avec le kit *S.c.* EasyCompTM Transformation kit, commercialisé par Invitrogen. A 50µl de cellules compétentes Y2HGold sont additionnés 250ng de vecteur recombinant pGBKT7 et 500µl de solution III. Le mélange est vortexé vigoureusement pendant 30 s puis incubé pendant 1h à 30°C avec un vortexage de 30 s tous les quarts d'heure. Pour finir, 400 et 100µl de cellules Y2HGold transformées sont étalés sur des boîtes de SD/-T. Les boîtes sont placées 3 jours à 30°C.

Pour la technique de double-hybride, il est impératif de confirmer au préalable d'un crible que la protéine appât n'active pas de manière autonome les gènes rapporteurs dans les cellules Y2HGold, en absence d'une protéine proie (autoactivation). Afin de s'en assurer, plusieurs clones indépendants préalablement obtenus sur les boîtes SD/-T sont striés sur des boîtes de milieu SD/-T/X- α -Gal et SD/-T/X- α -Gal/AbA. Les boîtes sont placées 3 jours à 30°C. Si la protéine appât n'autoactive pas les gènes rapporteurs, des colonies bleues blanches apparaissent sur le milieu SD/-T/X- α -Gal et aucune colonie ne se développe sur le milieu SD/-T/X- α -Gal/AbA. Dans le cas où la protéine active de manière autonome les gènes rapporteurs dans les cellules Y2HGold, des colonies bleues apparaissent sur les boîtes SD/-T/X- α -Gal et SD/-T/X- α -Gal/AbA.

Rencontre de la souche MAT α Y187 renfermant la banque ADNc et de la souche MAT α Y2HGold transformée pour la protéine fusion appât (criblage)

Une colonie Y2HGold de 2-3 mm contenant la protéine appât n'autoactivant pas, pGBKT7-TaGW2(1-320) dans ce travail de thèse, est inoculée dans 50 ml de milieu liquide SD/-T et incubée 22h, à 30h, sous agitation 250 rpm, afin d'atteindre une DO à 600 nm de 0,8. La suspension de levure est alors centrifugée à 1000g, pendant 5 min ; le surnageant est éliminé et le culot est finalement repris dans 4 ml de milieu liquide SD/-T afin d'atteindre une

Forme « tête de
mickey »



Forme feuille de
trèfle

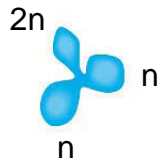


Figure 59 : Zygote de *S. cerevisiae* en formation. Les zygotes en formation ont une structure trilobée : les lobes du zygote correspondent aux deux cellules haploïdes parentales et au bourgeon diploïde. Certains zygotes ont la forme d'une feuille de trèfle, d'autres la forme d'une « tête de Mickey ».

densité cellulaire minimale de 1.10^8 par ml (la densité cellulaire est alors estimée à l'aide d'une cellule de comptage de Fuchs-Rosenthal). Un millilitre de la banque ADNc dans la souche Y187 est décongelé, mis en contact des 4ml de la souche Y2HGold-pGBKT7-TaGW2(1-320) dans un erlenmeyer stérile de 2L et additionné de 45 ml de milieu YPAD liquide 2X. Le tube contenant précédemment la banque Y187 est rincé 2 fois avec 1 ml de YPAD 2X et les 2 ml récupérés sont ajoutés à l'erlenmeyer de 2L. Le tout est incubé pendant 21h30, à 30°C, sous agitation 30-50 rpm. Le début d'apparition des zygotes en forme de « tête de Mickey » est alors vérifié avec un microscope à contraste de phase (Figure 59). La suspension de levure est centrifugée à 1000g pendant 10 min. L'erlenmeyer de 2L est lavé 2 fois avec 50 ml de milieu liquide YPAD 0,5X additionné de 50µg/ml de kanamycine. Le surnageant est éliminé délicatement et le culot est resuspendu dans les 100 ml de YPAD 0,5X de rinçage de l'erlenmeyer. La suspension de levure est de nouveau centrifugée à 1000g pendant 10 min et le culot est repris dans 10 ml de 0,5X YPAD/kanamycine. Le volume final de la suspension de cellules reprises est mesuré. Afin d'estimer le nombre total de clones diploïdes dans cette rencontre, 100 µl de la suspension de levures diploïdes diluée à $1/10^{\text{ème}}$, $1/100^{\text{ème}}$, $1/1000^{\text{ème}}$ et $1/10000^{\text{ème}}$ sont étalés sur boîtes SD/-T, SD/-L et SD/-L/-T. Les boîtes sont incubées pendant 3 à 5 jours, à 30°C. Le calcul du nombre total de clones diploïdes se fait comme suit :

Nombre de clones diploïdes = Nombre de colonies sur la boîte SD/-L/-T dilution $1/1000^{\text{ème}}$ x volume final de la suspension de cellules reprises (en ml) x 10 x 1000

Il est impératif qu'au minimum un million de cellules de levure diploïdes soit criblé au risque de ne pas détecter ensuite de vraies interactions sur le milieu de sélection SD/-L/-T/X- α -Gal/Aba.

Afin de faciliter l'étalement, le reste de la suspension de cellules de levures diploïdes est dilué par un ajout de 5 ml de 0,5X YPAD/kanamycine (50µg/ml) et étalé sur des boîtes SD/-L/-T/X- α -Gal/Aba de 15 cm de diamètre à raison de 200µl par boîte. Les boîtes sont incubées 3 à 5 jours, à 30°C.

Les clones bleus sélectionnés sur ce milieu sont des clones diploïdes présentant une interaction positive. Une sélection sur ce milieu sans leucine ni tryptophane permet une sélection de cellules de levure possédant à la fois au moins en un exemplaire le vecteur pGBKT7-protéine appât et en un exemplaire un vecteur recombinant pGADT7-ADNc. Les

clones présentant une interaction entre la protéine appât d'intérêt et une protéine proie de la banque peuvent pousser sur le milieu avec aureobasidine A et sont bleu, du fait de l'activation de l'expression des gènes rapporteurs *AURI-C* et *MEL1* respectivement.

Toutes les colonies bleues sont individuellement striées sur des boîtes du milieu de plus forte stringence SD/-L/-T/-H/-Ad/Aba. Les boîtes sont incubées 3 à 5 jours, à 30°C. Par cette opération, seuls les clones présentant une activation des 3 gènes rapporteurs (*HIS*, *ADE* et *AURI-C*) sont sélectionnés. Cette opération est renouvelée pour les clones ayant poussé à l'issue de cette première sélection.

Confirmation et identification des protéines interactrices

Choix des clones ne présentant qu'un seul vecteur recombinant pGADT7-proie et ségrégation de ce plasmide.

Afin de ne sélectionner dans un premier temps que des clones ne possédant qu'un seul vecteur proie, une PCR sur colonies de levures est réalisée à l'aide du couple d'amorce « pGAD-insert screening », avec le mix réactionnel EconoTaq Plus Green (Lucigen), dans un volume réactionnel de 40 µl, selon les recommandations du fournisseur. A cette étape, il est très important de prendre très peu de cellules pour la PCR, un excès de cellules inhibant la réaction PCR. Les amorces du couple « pGAD-insert screening » permettent une amplification de l'insert ADNc contenu dans le vecteur pGADT7-Rec. Cette réaction PCR est effectuée sur chacun des clones ayant poussé à l'issue des deux cycles de sélection sur milieu SD/-L/-T/-H/-Ad/Aba. Quinze microlitres de chaque réaction PCR sont déposés et analysés sur un gel d'agarose à 1% afin de déterminer le nombre de produits d'amplification de taille différente dans le clone testé. La présence de plus d'un produit d'amplification traduit la présence de plus d'un plasmide pGADT7-Rec contenant un ADNc proie. En effet, lors de la création de la banque ADNc, les cellules de levures Y187 peuvent être transformées avec plus d'un vecteur recombinant pGADT7-proie.

Les produits PCR des clones présentant un seul produit d'amplification sont séquencés (Beckman Coulter Genomics) à l'aide de l'amorce universelle T7, afin de connaître l'identité de l'ADNc proie.

Les colonies présentant plusieurs plasmides recombinants sont striées plusieurs fois de suite sur des boîtes de milieu SD/-L/-T/-H/-Ad/Aba, en ne prenant qu'un seul clone à chaque fois,

afin de ségréger les plasmides contenant l'ADNc proie interacteur. Les clones sont de nouveau testés par PCR sur colonie de levures afin de déterminer s'il ne reste qu'un seul plasmide recombinant pGADT7-Rec dans le clone, et sont finalement séquencés à leur tour.

Récupération du plasmide recombinant pGADT7-proie.

Le séquençage peut avoir mis en évidence des réplicats d'une même séquence ADNc. Un seul représentant est choisi pour chaque séquence et une colonie de celui-ci est striée dans un carré de 1cmx1cm, sur une boîte SD/-L. Les boîtes sont alors incubées 3 jours à 30°C. Les plasmides de chacune des colonies isolées sont extraits à l'aide du kit « Easy Yeast Plasmid isolation kit » (Clontech), selon les recommandations du fournisseur. Cinquante microlitres de souche DH5 α d'*E.coli* chimiquement compétentes (Invitrogen) sont transformées avec 3 μ l d'extraction plasmidique, selon les recommandations du fournisseur. Les bactéries DH5 α transformées sont étalées sur des boîtes de LB supplémentées en ampicilline (100 μ g/ml) afin de sélectionner la présence du vecteur recombinant pGADT7-proie. Les plasmides recombinants sont séquencés (Beckman Coulter Genomics) avec l'amorce universelle T7 sens.

Première élimination des faux-positifs par le test 2 à 2.

Les colonies positives obtenues à l'issue des sélection successives sur le milieu SD/-L/-T/-H/-Ad/Aba peuvent correspondre à une vraie interaction physique entre la protéine appât et une protéine proie, mais peuvent également être des colonies de levures « faux-positives » possédant une protéine proie capable d'activer les gènes rapporteurs sans interagir physiquement avec la protéine fusion DBD-appât. Ces protéines proie générant des interactions « faux-positives » peuvent être des protéines affectant la structure de la chromatine ou des protéines ayant une faible affinité ou une affinité aspécifique pour les régions promotrices des gènes rapporteurs.

Ainsi, la capacité de la protéine fusion AD-proie à activer les 4 gènes rapporteurs en l'absence de la protéine fusion DBD-appât est testée dans la souche Y2HGold. Cinquante microlitres de cellules de levures Y2HGold compétentes sont co-transformées d'une part avec 250 ng de chacun des vecteurs recombinants pGBKT7-appât et 250 ng pGADT7-proie, et d'autre part avec 250 ng de vecteur pGBKT7 vide et de 250 ng de vecteur recombinant pGADT7-proie, à l'aide du kit *S.c.* EasyCompTM Transformation (Invitrogen), selon les recommandations du fournisseur. Les cellules Y2HGold co-transformées sont tout d'abord étalées sur des boîtes de milieu SD/-L/-T et incubées pendant 3 jours à 30°C, puis striées sur un milieu SD/-L/-T/-H/-

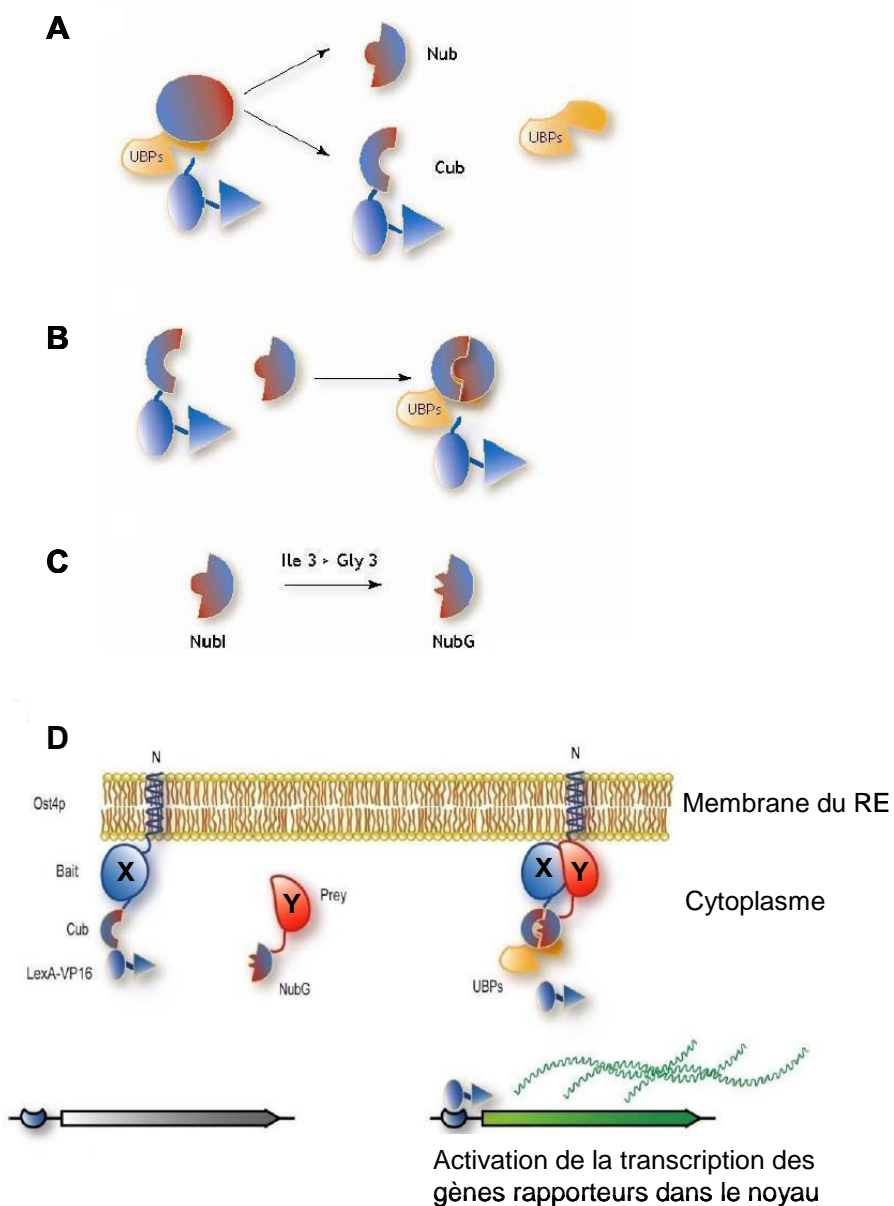


Figure 60 : Principe du système double-hybride DUALHunter, basé sur la modularité de l'ubiquitine (split-ubiquitine). (A) L'ubiquitine peut être séparée en deux parties Nub et Cub. (B) Lorsque Nub et Cub sont co-exprimées dans une même cellule, elles reforment spontanément une ubiquitine (split-ubiquitine) qui est reconnue par l'enzyme UBP. (C) La mutation de l'isoleucine 3 dans Nub (NubG) abolit la forte affinité entre Nub et Cub. (D) Dans le système DUALHunter, une protéine appât X est fusionnée en C-terminale de la protéine Ost4p (pour la rétention de la protéine appât dans le cytoplasme, sur la face externe de la membrane du RE) et en N-terminale de Cub-LexA-VP16. Une protéine proie Y est fusionnée en C-terminale de NubG. L'interaction entre les protéines X et Y autorise la reconstitution de la split-ubiquitine, qui est reconnue par l'enzyme UBP. L'enzyme UBP clive la liaison peptidique après Cub, ce qui libère le facteur de transcription LexA-VP16, qui peut activer l'expression des gènes rapporteurs dans le noyau.

Ad/Aba et incubées pendant 3 jours à 30°C. Une protéine proie est considérée « faux-positif » si elle active l'expression des gènes rapporteurs quelque soit la présence ou l'absence de la protéine fusion appât.

12.2 Le double-hybride cytoplasmique

Le système double-hybride cytoplasmique DUALHunter (Dualsystems Biotech) repose sur la possibilité de séparation en deux de l'ubiquitine puis sa reconstitution (split-ubiquitine). L'avantage majeur de ce système double-hybride cytoplasmique est qu'il permet la recherche d'IPP par technique de double-hybride pour des protéines fortement autoactivatrices dans un système double-hybride nucléaire, comme c'est le cas des facteurs de transcription par exemple, ou de protéines avec des domaines fortement acides. En effet, contrairement au double-hybride classique, les interactions ne sont pas forcées à avoir lieu dans le noyau pour l'activation des gènes rapporteurs. Par conséquent des protéines avec une affinité aspécifique pour la chromatine ou pour la machinerie générale de transcription n'activent pas de manière autonome l'expression des gènes rapporteurs, situés eux dans le noyau.

Le système split-ubiquitine a été inventé par Nils Johnsson et Alexander Varshavsky (1994), qui montrèrent que l'ubiquitine pouvaient être séparée en deux moitiés, appelés Nub (ubiquitine N-terminale) et Cub (ubiquitin C-terminale ; Figure 60, A). Lorsque les moitiés sont exprimées séparément dans une cellule, elles adoptent un repliement partiel de l'ubiquitine et ne peuvent être reconnues par des enzymes UBP (ubiquitin protease), qui reconnaissent la conformation tridimensionnelle native de l'ubiquitine. Si Nub et Cub sont co-exprimées dans une même cellule, leur forte affinité l'une pour l'autre amène à la reconstitution spontanée d'une ubiquitine (split-ubiquitine), qui présente une conformation tridimensionnelle très proche de l'ubiquitine native et est reconnue par les enzymes UBP (Figure 60, B). La forte affinité naturelle entre Nub (aussi appelée NubI) et Cub peut être abolie par la substitution de l'isoleucine en position 3 de NubI en glycine (Figure 60, C). La moitié Nub mutée est ainsi appelée NubG et ne présente presque plus spontanément d'affinité pour Cub lorsque les deux moitiés sont exprimées dans une même cellule.

Dans un système split-ubiquitine, une protéine X est fusionnée à NubG et une protéine Y à Cub. L'interaction entre les deux protéines X et Y rapprochent dans l'espace les deux moitiés

d'ubiquitine, reformant une split-ubiquitine qui est alors reconnue et clivée par les UBP. Dans le système DUALHunter, la protéine appât est une protéine fusion Ost4-appât-Cub-LexA-VP16. La protéine du réticulum endoplasmique de la levure Ost4 permet un ancrage membranaire de la protéine appât, pour lui éviter d'être envoyée spontanément dans le noyau (dans le cas d'une protéine nucléaire comme un facteur de transcription par exemple). LexA-VP16 est un facteur de transcription. La protéine proie est fusionnée à NubG. L'interaction entre les protéines appât et proie permet la reconstitution de la split-ubiquitine, qui est reconnue par les enzymes UBP, qui clivent la chaîne peptidique entre Cub et LexA-VP16 (Figure 60, D). Le facteur de transcription LexA-VP16 est alors libéré et est envoyé spontanément dans le noyau, où il va activer l'expression de gènes rapporteurs.

Construction de la banque ADNc

La banque ADNc normalisée a été construite à l'aide du kit EasyClone normalized cDNA library construction (Dualsystems Biotech). Seuls quelques détails propres à nos expérimentations seront mentionnés ci-après. Des grains récoltés à 80, 160 et 220°J ont été rassemblés et broyés ensemble à l'azote liquide dans un mortier préalablement refroidi à l'azote liquide. Les ARN totaux ont été extraits par une extraction ARN au tampon Borate comme décrit p.X.

Brièvement, la synthèse d'ADNc simple brin a été réalisée à l'aide de 2µg d'ARN totaux, et d'amorces-adaptateurs CDS-3M et PlugOligo-3M, selon les recommandations du fournisseur. La réaction de transcription inverse a été réalisée en présence d'inhibiteur de RNase (20U) dans le milieu réactionnel (RNaseOUT, Invitrogen). La synthèse d'ADNc double brin a été entreprise par une amplification sur 21 cycles ; puis les produits d'amplification ont été purifiés sur colonne à l'aide du kit Nucleospin Extract II (Macherey-Nagel), selon les recommandations du fournisseur avec quelques modifications. En effet, l'élution des produits purifiés a été réalisée en deux temps : après un chauffage des colonnes à 70°C, dans un bain-sec, pendant 3 min ; 20µl de tampon d'élution ont été ajoutés sur la colonne, puis la colonne a été placée à 70°C, dans un bain-sec, pendant 5 min, centrifugée 1 min à 40g puis 1 min à 11000g. L'éluat a été réservé et cette étape a été renouvelée avec 20 µl de tampon d'élution. Les éluats ont été poolés et l'ensemble a été dosé sur gel d'agarose à l'aide d'une échelle de poids moléculaire de quantification. Un microgramme deux cent d'ADNc double brin ont ensuite été normalisés selon les recommandations du fournisseur. La

totalité des produits d'amplification des ADNc double brin normalisés a été purifié sur colonnes à l'aide du kit Nucleospin Extract II (Macherey-Nagel), selon les recommandations du fournisseur avec les modifications précédentes pour l'étape d'élution (élution en 2 temps). La population ADNc double brin normalisée finale ainsi obtenue est flanquée par des adaptateurs possédant les sites Sfi nécessaires au clonage ultérieur de la population ADNc dans le vecteur de la banque pPR3-N.

Deux microgrammes de vecteur pPR3-N et 2µg d'ADNc double brin normalisés sont digérés par 20 unités d'enzyme SfiI (Promega), pendant 4h, à 50°C. Les produits de digestion sont ensuite déposés sur gel d'agarose puis purifiés à l'aide du kit Nucleospin Extract II (Macherey-Nagel), selon les recommandations du fournisseur avec les modifications précédentes pour l'étape d'élution (élution en 2 temps). Le clonage directionnel des ADNc double brin dans le vecteur pPR3-N linéarisé repose sur la ligation des extrémités cohésives et complémentaires Sfi par l'enzyme T4 DNA ligase. La réaction de ligation a été entreprise selon les recommandations du fournisseur, avec 3 unités d'enzyme T4 DNA ligase, dans un thermocycleur. Avant transformation d'*E.coli*, le produit de ligation a été précipité à l'éthanol selon le protocole mentionné en annexe 5. Afin d'amplifier la banque ADNc, le produit de ligation ainsi traité a été transformé dans la souche DH10B d'*E.coli* électrocompétente (Invitrogen), à l'aide de l'électroporateur *E.coli* pulser (BIORAD) selon les recommandations des deux fournisseurs. Les bactéries transformées ont été étalées sur des boîtes de milieu LB supplémenté en ampicilline (100µg/ml) pour la sélection du plasmide pPR3-N, incubées 1 nuit à 37°C, puis les bactéries ont été récoltées selon les recommandations du manuel du kit Easyclone (Dualsystems Biotech). La banque plasmidique ADNc a finalement été extraite des bactéries par maxi-préparation, à l'aide du kit NucleoBond Xtra Maxi Plus (Macherey-Nagel).

Construction de la protéine fusion appât

Avant le clonage de la protéine d'intérêt dans le vecteur pDHB1, il est important de s'assurer que la séquence d'intérêt ne possède pas de sites de reconnaissance de l'enzyme de restriction Sfi. L'ADNc pleine-longueur de *TaGW2-A* sans le codon stop terminal a été amplifié par PCR sur ADN plasmidique (p221-TaGW2-A), avec la Taq Advantage (Clontech) selon les recommandations du fournisseur, à l'aide du couple d'amorces SfiI-TaGW2. Les amorces du couple SfiI-TaGW2 sont constituées (de 5' vers 3') de 8 nt quelconques, puis de 13 à 15 nt correspondant aux sites SfiA et B (sites de reconnaissance de l'enzyme Sfi) et enfin

d'une vingtaine de nt spécifiques de la séquence ADNc de *TaGW2-A* à amplifier. Le produit d'amplification a été déposé sur gel d'agarose puis purifié sur colonne à l'aide du kit Nucleospin Extract II (Macherey-Nagel), selon les recommandations du fournisseur. Le produit PCR purifié a alors été dosé sur gel d'agarose avec une échelle de poids moléculaire de quantification. Un et quatre microgrammes de produits PCR purifiés et de vecteur pDHB1 ont été digérés par 20 unités d'enzyme SfiI (Promega), dans un mélange réactionnel décrit par le fournisseur, pendant 3h30 à 50°C (bain-marie). Les produits de digestion ont finalement été purifiés sur colonne à l'aide du kit Nucleospin Extract II (Macherey-Nagel), selon les recommandations du fournisseur avec les modifications précédentes pour l'étape d'élution (élution en 2 temps). La réaction de ligation entre le vecteur pDHB1 linéarisé et le produit PCR digérés par SfiI puis purifiés a été réalisée à l'aide de l'enzyme T4 DNA ligase (Promega), selon les recommandations du fournisseur, en respectant un rapport insert/vecteur de 3 :1. Des bactéries *E.coli* de souche DH5 α chimiquement compétentes (Invitrogen) ont finalement été transformées, étalées sur des boîtes de milieu LB supplémenté en kanamycine (50 μ g/ml) et incubées une nuit à 37°C. La construction a été vérifiée par séquençage (Beckman Coulter Genomics).

La souche de levure NMY51 a été rendue compétente puis a été transformée avec la construction pDHB1-TaGW2 à l'aide du kit DSY Yeast Transformation (Dualsystems Biotech), selon les recommandations du fournisseur. La souche NMY51 transformée a finalement été étalée sur boîte SD/-L et incubée 3 jours, à 30°C. Des stocks glycérol de cette souche NMY51 ont été faits.

Les tests fonctionnels n°1, 2 et 3 (respectivement tests de l'expression et d'auto-activation de la protéine chimère, et choix du milieu sélectif pour le criblage de la banque) ont été réalisés comme mentionné dans le manuel de l'utilisateur du DUALHunter starter kit (Dualsystems Biotech).

Criblage de la banque ADNc avec la construction appât

La souche transformée NMY51 pDHB1-TaGW2 a été fraîchement striée sur des boîtes de milieu SD/-L, afin d'assurer un taux de transformation ultérieur élevé de la souche, puis transformée avec 7 μ g d'ADN de la banque ADNc plasmidique selon les recommandations du manuel de l'utilisateur du DUALHunter starter kit (Dualsystems Biotech). Des aliquots des

levures NMY51 transformées ont été étalés sur des boîtes de milieu SD/-L/-T pour un titrage de la rencontre, et le restant sur des boîtes de milieu sélectif pour les interacteurs positifs, déterminé au préalable au cours de l'étude pilote (test n°3). Ce milieu restrictif sélectionne pour l'expression des gènes rapporteurs *HIS3* et *ADE2*.

Les colonies de levures obtenues sur le milieu restrictif pour le crible sont ensuite striées sur des boîtes de ce même milieu frais. L'activation de l'expression du gène rapporteur LacZ est alors testée pour ces clones de levures positifs à l'aide du kit HTX (High Throughput β -galactosidase assay ; Dualsystems Biotech), selon les recommandations du fournisseur.

Confirmation et identification des protéines interactrices

Une PCR sur colonie de levures est réalisée sur chacun des clones positifs, à l'aide du couple d'amorce pPR3-N et du mélange réactionnel EconoTaq (Lucigen), dans un volume réactionnel de 40 μ l. Quinze microlitres de chaque réaction PCR sont déposés et analysés sur un gel d'agarose à 1% afin de déterminer le nombre potentiel de vecteur proie dans le clone positif testé. Dans un premier temps, seuls les produits PCR des clones présentant un seul produit d'amplification sont séquencés (Beckman Coulter Genomics) afin de connaître l'identité de l'ADNc proie.

Le séquençage peut avoir mis en évidence des réplicats d'une même séquence ADNc, et un seul représentant est choisi pour chaque séquence pour la suite des étapes. Le plasmide proie est extrait des levures à l'aide du kit PureLink QuickPlasmid (Invitrogen) et de microbilles de verres lavées à l'acide (Sigma-Aldrich). Le plasmide proie extrait est finalement entreposé dans des bactéries DH5 α (Invitrogen). Les bactéries transformées sont étalées sur des boîtes de milieu LB supplémenté en ampicilline (100 μ g/ml) pour la sélection du plasmide pPR3-N.

L'élimination des clones positifs est finalement réalisée par un test 2 à 2 dans la souche de levure NMY51, selon les recommandations du manuel de l'utilisateur du DUALHunter starter kit (Dualsystems Biotech), avec quelques modifications : en plus d'être testée contre le vecteur pDHB1 vide, la protéine proie est également testée pour une interaction avec une protéine appât différente de la protéine appât utilisée pour le crible. Seules les protéines proie capables d'interagir uniquement avec la protéine appât utilisée pour le crible, au début de l'analyse, sont conservées et peuvent être considérées comme des interacteurs potentiels de la protéine appât.

Chapitre I :
Caractérisation moléculaire et biochimique de
TaGW2

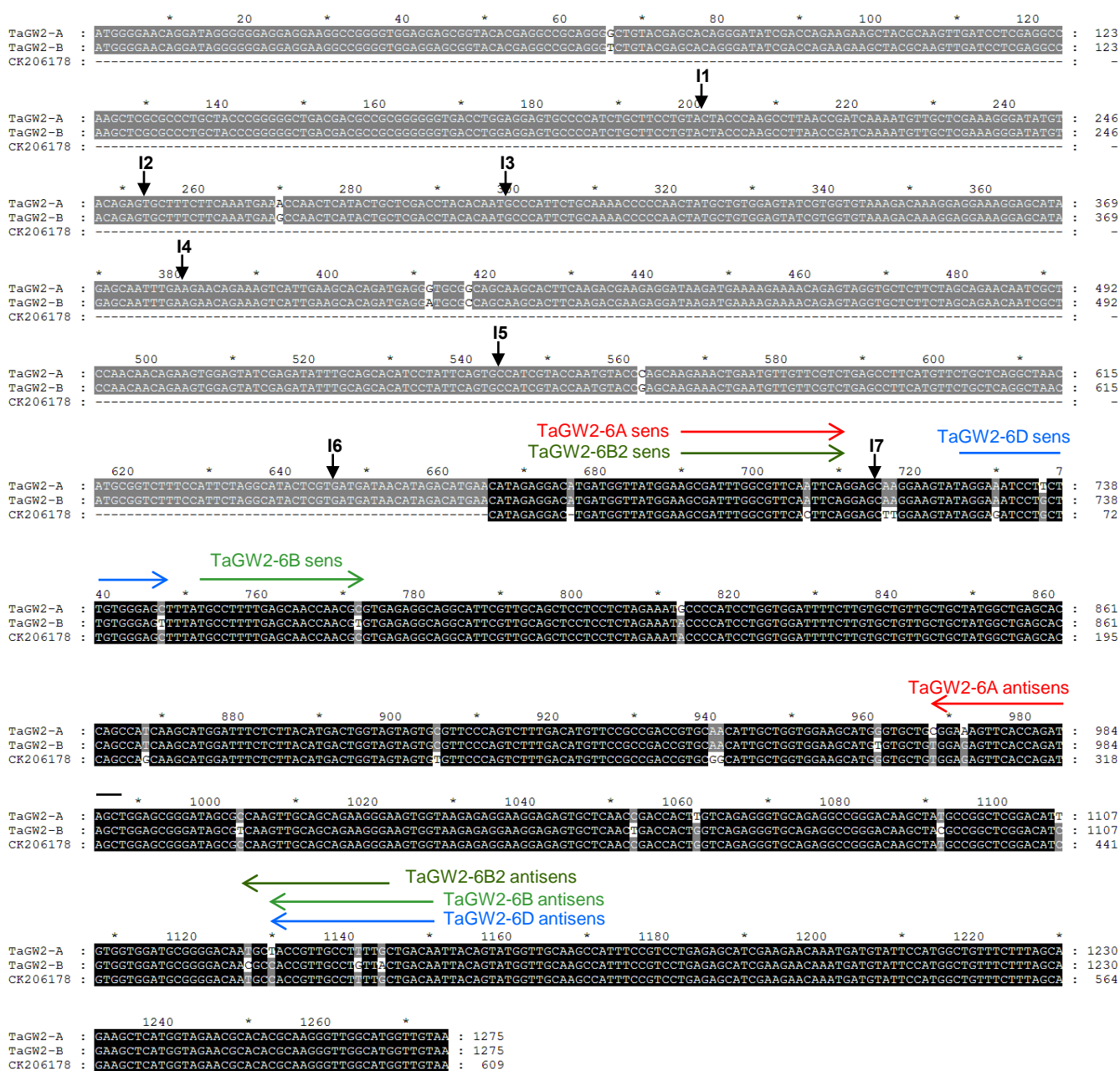


Figure 61: Alignement des séquences nucléotides ADNc des homologues *TaGW2-A* et *-B* de *T. aestivum* cv Récital, et de l'EST CK206178 de *T. aestivum* cv Chinese Spring. Les nucléotides identiques entre 2 et 3 séquences sont colorés en gris et noir, respectivement. Les nombres indiquent les positions des nucléotides dans les séquences. Les positions des introns (I1 à I7) ainsi que les sites d'hybridation des différents couples d'amorces utilisés pour l'assignation sur les lignées aneuploïdes (couples d'amorces TaGW2-6A, -6B et -6D) et le développement de marqueurs polymorphes (couples d'amorces copies-spécifiques TaGW2-6A, -6B2 et -6D) sont indiqués au dessus des séquences.

1. Obtention des séquences codantes des 3 homéologues de *TaGW2*

Afin d'obtenir les séquences complètes codantes (CDS) des trois homéologues *TaGW2-A*, *TaGW2-B* et *TaGW2-D*, nous avons dessiné un couple d'amorce à partir de la séquence CDS de *GW2* chez le riz (EF447275.1) et amplifié par PCR semi-quantitative la population ADNc pleine-longueur, toutes copies confondues, de *TaGW2* dans des feuilles de *T. aestivum* cv Récital. L'analyse des séquences amplifiées a révélé deux séquences différentes, probablement celles de deux homéologues de *TaGW2*. Suite à une analyse des séquences EST de blé tendre disponibles dans Genbank, nous avons identifié deux groupes d'EST correspondant à chacune des séquences ADNc obtenues par amplification, et une séquence EST (CK206178) présentant quelques différences avec les deux groupes d'EST identifiés. Il a alors été émis l'hypothèse que l'EST CK206178 représentait le troisième homéologue de *TaGW2*. Des couples d'amorces spécifiques de chaque groupe de séquences (couples d'amorces *TaGW2-6A*, *TaGW2-6B*, *TaGW2-6D*) ont été dessinés et ont permis l'assignation, par PCR sur des lignées aneuploïdes de *T.aestivum* cv Chinese Spring et des lignées de délétion, de chaque groupe de séquence au bras court des chromosomes 6A (bin de délétion 6AS1-0.35), 6B (bin de délétion 6BS5-0.76), et 6D (bin de délétion 6DS2-0.45).

Les résultats ont révélé que les séquences complètes codantes de *TaGW2-A* et *-B* avaient été obtenues, et que l'EST CK206178 correspondait à la séquence partielle en 3' de *TaGW2-D* (des nucléotides 667 à 1275 des séquences ADNc de *TaGW2-A* et *-B* ; Figure 61). Les séquences ADNc de *TaGW2-A* et *-B* ont été déposées dans Genbank et possèdent respectivement les numéros d'accessions suivants : JN896622 et JN896623.

Entre-temps, une nouvelle tentative d'obtention de l'ADNc complet de *TaGW2-D* a été entreprise par RT-PCR sur des grains de blé tendre au stade 600°CJours, à l'aide du couple d'amorces *TaGW2FL_CDS* dessiné sur la séquence ADNc de *TaGW2-A* et *-B*. L'analyse des séquences amplifiées a révélé l'obtention des trois CDS *TaGW2-A*, *-B* et *-D*. Les ADNc des homéologues *TaGW2* présentent une identité de séquence en nucléotide d'environ 98% entre eux (98% entre *TaGW2-A* et *-D*, et *TaGW2-A* et *-B* ; 99% entre *TaGW2-B* et *-D*). Vingt-quatre SNPs (Single nucleotide polymorphism) ont été identifiés entre les 3 homéologues (Figure 62, A). Les séquences primaires putatives des homéologues *TaGW2* montrent 97% d'identité de séquence en acides aminés entre eux. Le domaine RING, allant de l'acide aminé 62 à 104, est strictement conservé entre les protéines *TaGW2-A*, *-B* et *-D* prédites (Figure 62, B).

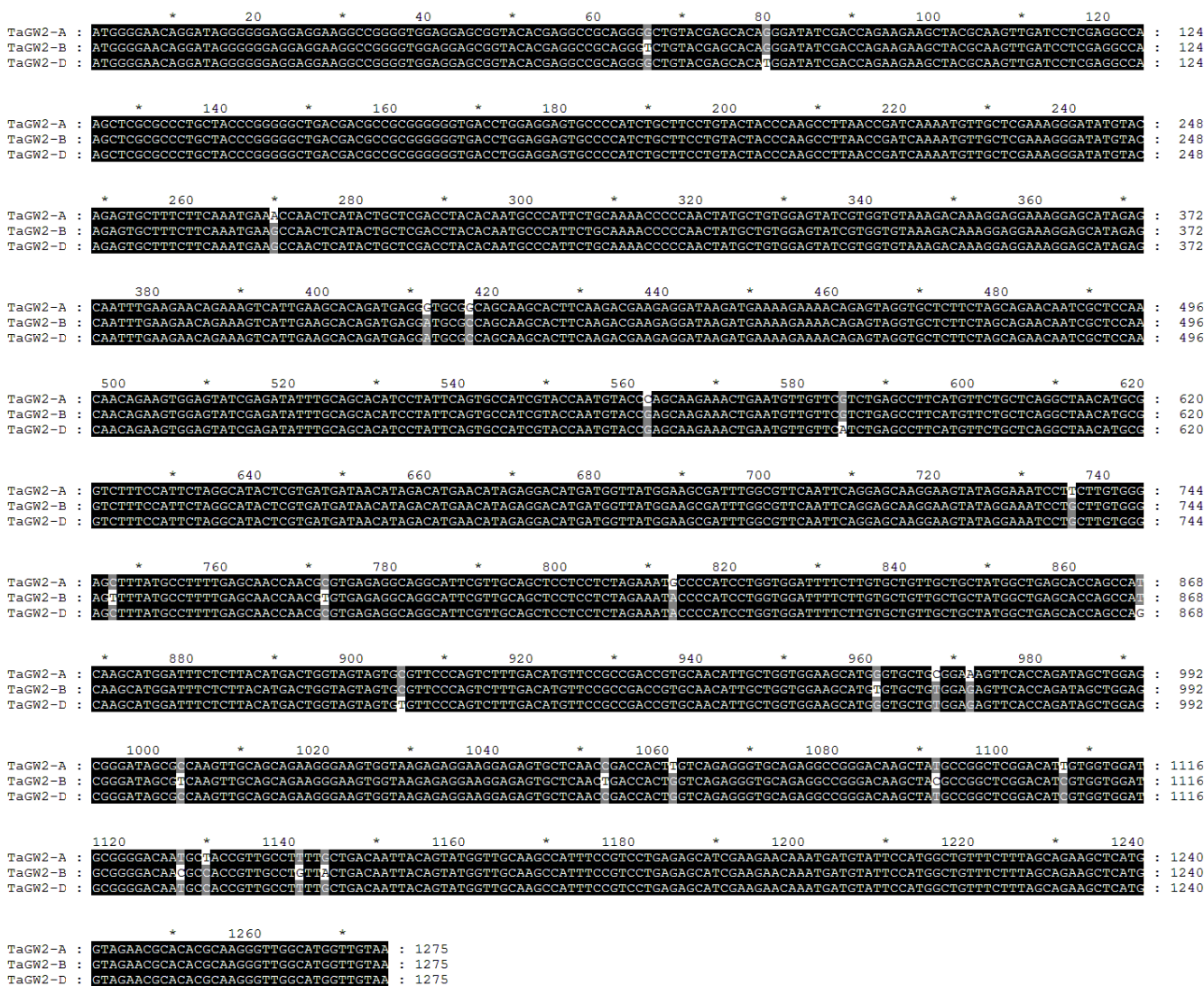
A**Domaine RING****B**

Figure 62 : Alignement des séquences complètes nucléotidiques ADNc (A) et des séquences complètes primaires protéiques (B) des homéologues TaGW2-A, -B et -D, de *T. aestivum* cv Récital. Les nucléotides (A) et acides aminés (B) identiques entre 2 et 3 séquences sont colorés en gris et noir, respectivement. Les nombres indiquent les positions des nucléotides dans les séquences. (B) Le domaine RING est encadré en rouge.

Tableau 7 : Origines et allèles au marqueur TaGW2_A_Y1061 des treize variétés de blé tendre utilisées pour la recherche de diversité nucléotidique dans les copies de *TaGW2*.

Accession	Origine	Allèle au marqueur TaGW2_A_Y1061
Apache	France	C
Arche	France	T
Balkan	Yougoslavie	T
Chinese Spring	Chine	C
Courtot	France	T
Glenlea	Canada	T
M708//G25/N163	Israël	C
Miskaagani	Liban	T
Opata85-Mex	Mexique	T
Ralet	France	T
Recital	France	T
Renan	France	T
Seu Seun 27	Corée	T

Une analyse blastx de l'ADNc de *TaGW2* dans NCBI révèle une identité de séquence en acides aminés de 87,5% avec GW2 (EF447275) chez le riz ; 93,6% avec la protéine YRG1 (ABY51682) chez l'orge ; 89,6% avec un homologue (XP003571977) chez *Brachypodium* ; 83,3% et 82% avec les homologues ZmGW2-CHR4 et ZmGW2-CHR5 (NP001144494 et NP1143691, respectivement) chez le maïs ; et 44,1% avec un homologue (NP_565180) chez *Arabidopsis*.

L'ensemble des résultats qui suivent ont été générés sur la base des ADNc obtenus au début de mon travail de thèse, à savoir uniquement avec la connaissance de la séquence complète codante de *TaGW2-A* et *TaGW2-B*, et l'information EST CK206178 codant la séquence 3' de *TaGW2-D*.

2. Caractérisation génétique de *TaGW2-A* et *-B*

2.1. Développement de marqueurs polymorphes pour *TaGW2-A* et *TaGW2-B*

Afin de développer des marqueurs polymorphes pour les 3 copies de *TaGW2*, une PCR sur ADN génomique a été réalisée sur 13 variétés représentatives de la variabilité du blé tendre (Tableau 7), à l'aide des couples d'amorces copies-spécifiques TaGW2-6A, -6B2 et 6D. L'amorce sens des couples TaGW2-6A et TaGW2-6B2 s'hybrident dans l'exon 7 de *TaGW2*; et l'amorce sens du couple TaGW2-6D dans l'exon 8 (Figure 61). Les amorces anti-sens des 3 couples d'hybrident dans l'exon 8 de *TaGW2*. Les couples TaGW2-6A et -6B2 incluent donc une amplification de l'intron 7 de *TaGW2*, tandis que le couple TaGW2-6D ne permet qu'une amplification de la séquence exonique 8.

2.1.1. Pour *TaGW2-A*

Le séquençage de 1318 pb de *TaGW2-A* dans les 13 variétés a révélé un SNP intronique (intron 7), qui est une substitution C/T, en position 1061 de la séquence amplifiée. Ce marqueur a été appelé TaGW2_A_Y1061. Pour le génotypage ultérieur de ce marqueur, des couples d'amorces spécifiques à la fois de *TaGW2-A* et de chacun des allèles du marqueur ont été dessinés (couple d'amorces A-SNPintron7). L'amorce sens des couples est commune aux deux couples d'amorces et amène la spécificité pour *TaGW2-A*, et les amorces anti-sens des couples assurent l'amplification allèle-spécifique.

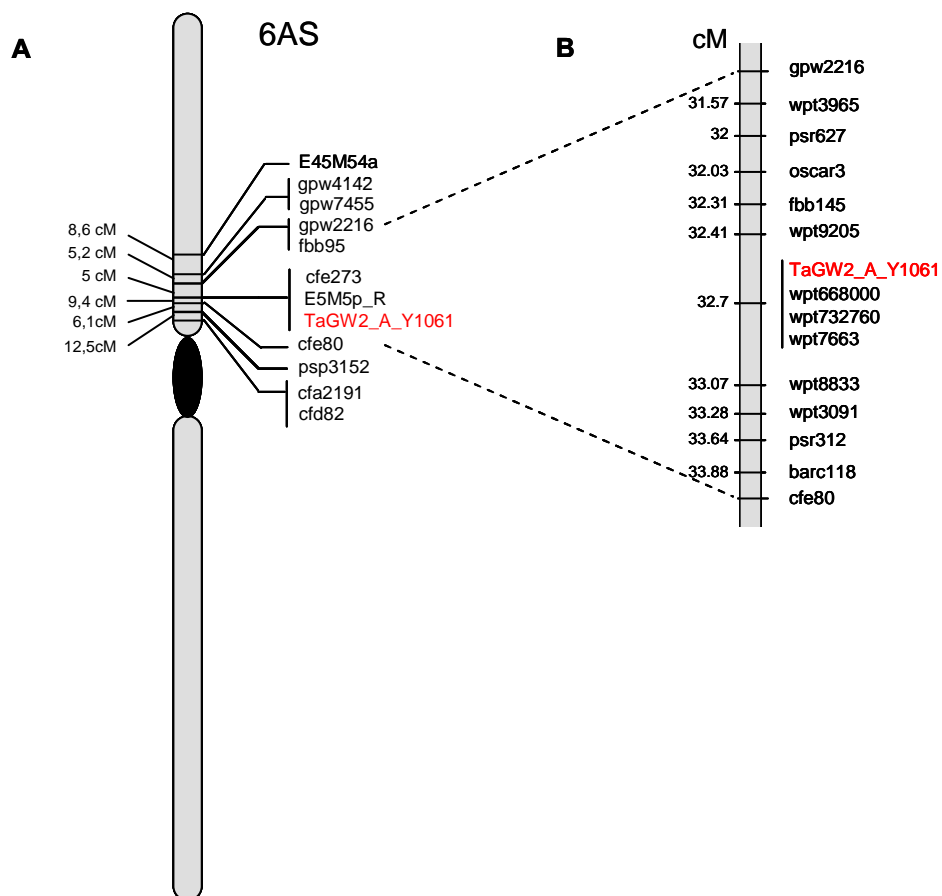


Figure 63 : Cartographie génétique de *TaGW2-A*, sur la population Ct x Cs. (A) Position du marqueur *TaGW2_A_Y1061* sur le 6AS, par étude de liaison génétique dans la population Ct x Cs. (B) Position génétique du marqueur *TaGW2_A_Y1061* sur la carte consensus.

2.1.2. Pour *TaGW2-B*

Dans les 1542 pb analysés pour *TaGW2-B*, trois SNPs et un polymorphisme d'insertion-délétion ont été identifiés dans l'intron 7. Les SNPs identifiés sont des substitutions T/G, G/A et C/G en position nucléotidique 447, 466 et 582 respectivement de la séquence amplifiée. Le polymorphisme d'insertion-délétion est en position 669 de la séquence amplifiée. L'ensemble de ces quatre polymorphismes ne donne que deux haplotypes, indiquant un taux élevé de déséquilibre de liaison. Un seul SNP a donc été considéré par la suite, le SNP T/G, et a été appelé *TaGW2_B_Y447*. Pour le génotypage ultérieur de ce marqueur, des couples d'amorces spécifiques à la fois de *TaGW2-B* et de chacun des allèles du marqueur ont été dessinés (B-SNP_{intron7}). Les amorces sens des couples assurent l'amplification allèle-spécifique, et l'amorce anti-sens des couples est commune et assure la spécificité vis-à-vis de *TaGW2-B*.

Le séquençage de 430 pb de *TaGW2-D* dans les 13 variétés n'a révélé aucun polymorphisme de séquence.

2.2. Cartographie génétique de *TaGW2-A* et *TaGW2-B*

2.2.1. Cartographie génétique de *TaGW2-A*

Le marqueur *TaGW2_A_Y1061* est polymorphe pour les variétés Courtot et Chinese Spring, possédant respectivement l'allèle T et C du marqueur. Une population d'haploïdes doublés générés par androgénèse possédant les variétés Courtot et Chinese Spring comme parent de population est disponible au sein de l'UMR 1095 GDEC, et le marqueur *TaGW2_A_Y1061* a donc été génotypé sur 169 lignées descendantes de cette population de cartographie Ct x Cs. L'analyse de liaison génétique entre le marqueur *TaGW2_A_Y1061* et l'ensemble des marqueurs de la population Ct x Cs (826 marqueurs en 2010) a été réalisée à l'aide du logiciel MAPMAKER/EXP 3.0. Le résultat de cartographie génétique de *TaGW2-A* obtenu est présenté en Figure 63, A. Le marqueur *TaGW2_A_Y1061* co-localise avec le marqueur AFLP E5M5p_R et le marqueur SSR dérivé d'EST cfe273 sur le chromosome 6AS, dans la population Ct x Cs. La cartographie de *TaGW2-A* a alors été définie sur une carte génétique consensus (Carte de référence Somers *et al*, 2004) afin d'enrichir les marqueurs moléculaires autour du marqueur *TaGW2_A_Y1061*. Le résultat de cartographie génétique du marqueur *TaGW2_A_Y1061* obtenu sur la carte génétique consensus est présenté en Figure

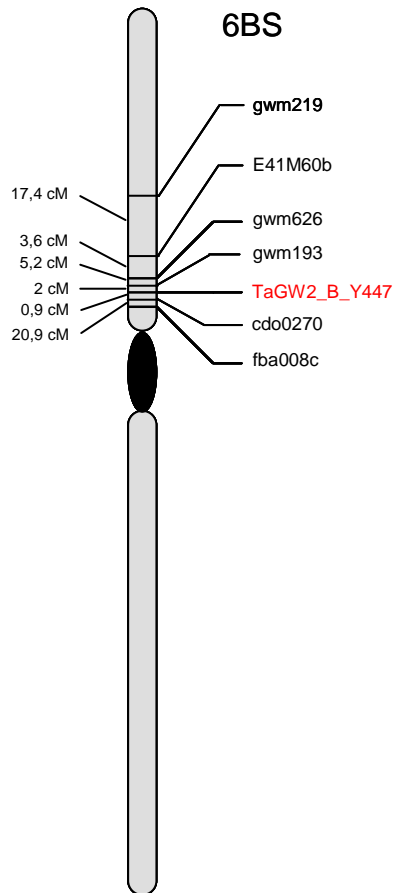


Figure 64 : Cartographie génétique de *TaGW2-B*, sur la population Re x R. Position du marqueur TaGW2_A_Y1061 sur le 6BS, par étude de liaison génétique dans la population Re x R.

63, B. Le marqueur TaGW2_A_Y1061 est situé en position 32,7 cM de la carte consensus du chromosome 6A, et co-localise avec les marqueurs wpt668000, wpt732760 et wpt7663.

Une recherche de QTL poids de 1000-grains a été entreprise sur l'ensemble des données de cette population Ct x Cs à l'aide du logiciel GenStat, mais aucun QTL pour le poids de 1000-grains n'a été détecté sur le chromosome 6AS.

2.2.2. Cartographie génétique de *TaGW2-B*

Le marqueur TaGW2_B_Y447 est polymorphe pour les variétés Renan et Récital, chacune possédant respectivement l'allèle T et G du marqueur. Le marqueur TaGW2_B_Y447 a été génotypé sur 194 lignées consanguines recombinantes (RIL) descendantes d'une population de cartographie issue du croisement entre les lignées Renan et Récital (Re x R). L'analyse de liaison génétique entre le marqueur TaGW2_B_Y447 et l'ensemble des marqueurs de la population Re x R (288 marqueurs en 2010) a été réalisée à l'aide du logiciel MAPMAKER/EXP 3.0. Le résultat de cartographie génétique de *TaGW2-B* obtenu est présenté en Figure 64. Le marqueur TaGW2_B_Y447 est localisé entre le marqueur microsatellite gwm193 et le marqueur RFLP cdo0270 sur le chromosome 6BS, à 2 et 0,9 cM de ces marqueurs respectivement, dans la population Re x R.

2.3. Analyse d'association dans la core collection mondiale de blé tendre (CC372)

Afin de regarder si du polymorphisme de séquences dans les copies A et B de *TaGW2* pouvait impacter des composantes du rendement chez le blé, nous avons entrepris une étude d'association à l'aide des marqueurs développés pour chacune des copies ainsi que le marqueur TaGW2_A_R(-593) développé par Su et ses collaborateurs (2011), sur une core-collection mondiale de blé tendre. L'analyse d'association a consisté à regarder s'il existait une relation entre les polymorphismes moléculaires identifiés dans *TaGW2-A* et *-B* et des variations phénotypiques pour les caractères mesurés sur cette population. Les marqueurs TaGW2_A_Y1061 et TaGW2_A_R(-593) pour la copie A de *TaGW2*, et le marqueur TaGW2_B_Y447 pour la copie B de *TaGW2* ont été génotypés sur les 372 individus de cette population par PCR sur ADN génomique. Le marqueur cfe273 co-localisant avec TaGW2_A_Y1061 dans la population Ct x Cs, il a également été inclus dans l'étude d'association. Le marqueur cfe273 avait été préalablement génotypé lors de l'établissement de cette core collection mondiale de blé tendre, et les données de génotypage ont été cordialement fournies par Jacques Bordes (UMR 1095 GDEC).

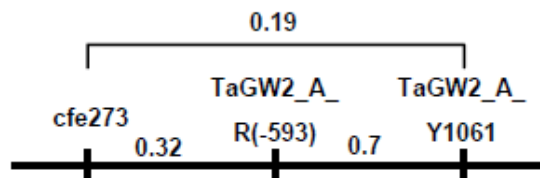


Figure 65 : Représentation schématique du déséquilibre de liaison observé entre les deux marqueurs pour la copie A de *TaGW2*, *TaGW2_A_R(-593)* et *TaGW2_A_Y1061*, et la marqueur *cfe273*, dans la core collection mondiale de blé tendre.

2.3.1. Résultats de l'analyse d'association pour *TaGW2-A*

Pour mémoire, le marqueur TaGW2_A_R(-593) identifié dans le promoteur de *TaGW2-A* par Su et ses collaborateurs (2011) est un marqueur SNP G/A. Afin de voir si ce marqueur était polymorphe dans le matériel utilisé pour l'analyse d'association dans cette étude, nous avons amplifié et séquencé une séquence de 949 pb de *TaGW2-A* dans 13 variétés de la core collection à l'aide du couple d'amorces TaGW2-6A-P1. L'analyse des séquences amplifiées a montré que le marqueur TaGW2_A_R(-593) est polymorphe dans ces variétés et qu'il peut donc être utilisé dans l'étude d'association. Su et ses collaborateurs avaient développé un marqueur CAPS pour le génotypage de ce marqueur. Cependant, afin d'uniformiser la méthode de génotypage au cours de ce travail de thèse, nous avons choisi de dessiner un nouveau couple d'amorces (couple d'amorces A-SNPR(-593)) permettant le génotypage de TaGW2_A_R(-593) par séquençage.

Le génotypage des marqueurs TaGW2_A_Y1061 et TaGW2_A_R(-593) a révélé un déséquilibre dans la proportion entre les allèles de chacun des marqueurs. Au marqueur TaGW2_A_Y1061, l'allèle T a été détecté dans 303 individus contre 46 individus portant l'allèle C du marqueur. Au marqueur TaGW2_A_R(-593), l'allèle G a été détecté dans 174 individus contre 69 portant l'allèle A. Un déséquilibre dans la proportion des allèles avait préalablement été identifié également au marqueur cfe273 (35 individus possédant l'allèle 303 et 318 individus possédant l'allèle 336). Quatre haplotypes de *TaGW2-A* figurent dans la population génotypée : les haplotypes G-T, A-T, G-C et A-C.

La recherche d'association entre le polymorphisme de séquence pour *TaGW2-A* et le marqueur cfe273, et la variation phénotypique pour 9 caractères a été entreprise à l'aide du logiciel TASSEL 2.0, en utilisant le modèle GLM (General Linear Model). Le déséquilibre de liaison (LD) a d'abord été calculé entre les 3 marqueurs avec la fonction 'Linkage disequilibrium' de TASSEL 2.0. Le plus haut LD a été observé entre les deux marqueurs pour TaGW2-A (LD de 0,71), tandis que le plus faible LD a été observé entre les marqueurs cfe273 et TaGW2_A_Y1061 (0,19). Le LD entre les marqueurs cfe273 et TaGW2_A_R(-593) était de 0,32. Sur la base de ces valeurs de LD, nous avons alors été capables de situer cfe273 par rapport aux 2 marqueurs pour *TaGW2-A* (Figure 65). Pour la recherche d'association entre les polymorphismes aux 3 marqueurs et les variations phénotypiques, nous avons pris en compte la structure génétique de la core collection comme co-variable dans le modèle GLM afin d'éviter au maximum les faux-positifs. Le résultat de l'analyse d'association est présenté dans le Tableau 8 :

Tableau 8 : Analyse d’association entre le polymorphisme aux marqueurs pour *TaGW2-A* et le marqueur *cfe273*, et la variation phénotypique de neuf caractères de rendement et de qualité du grain, dans les 372 individus de la core collection mondiale de blé tendre (CC372). Les valeurs mentionnées sont des moyennes des valeurs observées dans l’ensemble des variétés pour un allèle. Le nombre de lignées pour un allèle est donné entre parenthèses. Les résultats significatifs sont surlignés en jaune.

Marqueur	cfe273			TaGW2_A_R(-593)			TaGW2_A_Y1061		
Allele	303 (35)	336 (318)	p-value	G (174)	A (69)	p-value	T (303)	C (46)	p-value
Rendement (xx)	4,22	4,16	0,67	4,11	4,12	0,97	4,07	4,17	0,38
Longueur de l'épi (cm)	7,83	9,1	1,02E-04	9,21	8,55	0,01	9,11	8,63	0,1
Epillets par épi	20,06	20,64	0,2	20,67	20,46	0,57	20,66	20,49	0,67
Nombre de grains par épi	48,24	49,63	0,51	49,51	51,52	0,21	49,75	51,17	0,45
Taille du grain	4,92	5	0,07	5,01	4,98	0,35	5,03	5	0,51
Poids de 1000-grains (g)	37,95	41,4	7,40E-05	41,5	39,63	8,40E-03	41,51	41,04	0,56
Quantité d'azote par grain (mg d'azote par grain)	0,82	0,96	3,38E-07	0,97	0,86	3,38E-06	0,96	0,9	0,01
Teneur en protéines du grain (% de la masse sèche du grain)	12,44	13,29	1,80E-03	13,35	12,49	6,25E-05	13,27	12,5	1,70E-03
Dureté (xx)	24,91	36,37	0,015	37,35	23,74	2,63E-04	37,35	23,84	1,10E-03

Tableau 9 : Analyse d’ association entre le polymorphisme au marqueur pour *TaGW2-B* et la variation phénotypique de huit caractères de rendement et de qualité du grain, dans les 372 individus de la core collection mondiale de blé tendre (CC372). Les valeurs mentionnées sont des moyennes des valeurs observées dans l’ ensemble des variétés pour un allèle. Le nombre de lignées pour un allèle est donné entre parenthèses.

Marqueur	TaGW2_B_Y447		
Allele	T (145)	G (199)	p-value
Rendement (xx)	4,12	4,18	0,51
Longueur de l'épi (cm)	9,15	9,11	0,85
Epillets par épi	21,01	20,66	0,27
Nombre de grains par épi	50,04	49,83	0,89
Taille du grain	4,99	5,01	0,53
Poids de 1000-grains (g)	40,99	41,6	0,31
Teneur en protéines du grain (% de la masse sèche du grain)	13,49	13,29	0,28
Dureté (xx)	37,03	37,19	0,96

- Les 3 marqueurs sont significativement associés à une variation de la quantité d'azote dans le grain, à la teneur en protéines du grain et à la dureté du grain (toutes les p-values sont inférieures à 0,01). L'allèle 336 de cfe273, l'allèle G de TaGW2_A_R(-593) et l'allèle T de TaGW2_A_Y1061 sont associés à une augmentation en moyenne de :
 - ◊ 17%, 12,7% et 6,66% de la quantité d'azote par grain, respectivement
 - ◊ 6,83%, 6,88% et 6,16% de la teneur en protéines du grain, respectivement
 - ◊ 46%, 57,3% et 56,66% de la dureté du grain, respectivement.
- Les marqueurs cfe273 et TaGW2_A_R(-593) sont associés à une variation du poids de 1000-grains (p-values = $7,40 \times 10^{-5}$ and $8,40 \times 10^{-3}$, respectivement). L'allèle 336 de cfe273 et l'allèle G de TaGW2_A_R(-593) augmentent en moyenne le poids de 1000-grains de 9% et 4,71%, respectivement.
- Une association a également été observée entre une variation de la longueur de l'épi et du polymorphisme aux marqueurs cfe273 et TaGW2_A_R(-593) (p-values = $1,02 \times 10^{-4}$ et 0,01, respectivement) ; les allèles 336 et G de cfe273 et TaGW2_A_R(-593) sont associés à une augmentation de 16,2% et 7,71% de la longueur de l'épi.

2.3.2. Résultats de l'analyse d'association pour *TaGW2-B*

Le génotypage du marqueur TaGW2_B_Y447 a révélé un équilibre dans la proportion de chaque allèle du marqueur au sein des 372 individus de la core collection mondiale de blé tendre. L'allèle T a été détecté dans 145 individus et l'allèle G dans 199 individus (Tableau 9).

Tout comme précédemment, la recherche d'association entre le marqueur TaGW2_B_Y447 et la variation phénotypique pour 8 caractères a été entreprise à l'aide du logiciel TASSEL 2.0, en prenant en compte la structure génétique de la population comme co-variable dans le modèle GLM (General Linear Model). Le résultat de l'analyse d'association est présenté dans le Tableau 9 : aucune association n'a été retrouvée entre le polymorphisme identifié pour *TaGW2-B* et une variation pour l'un des 8 caractères phénotypés sur la core collection.

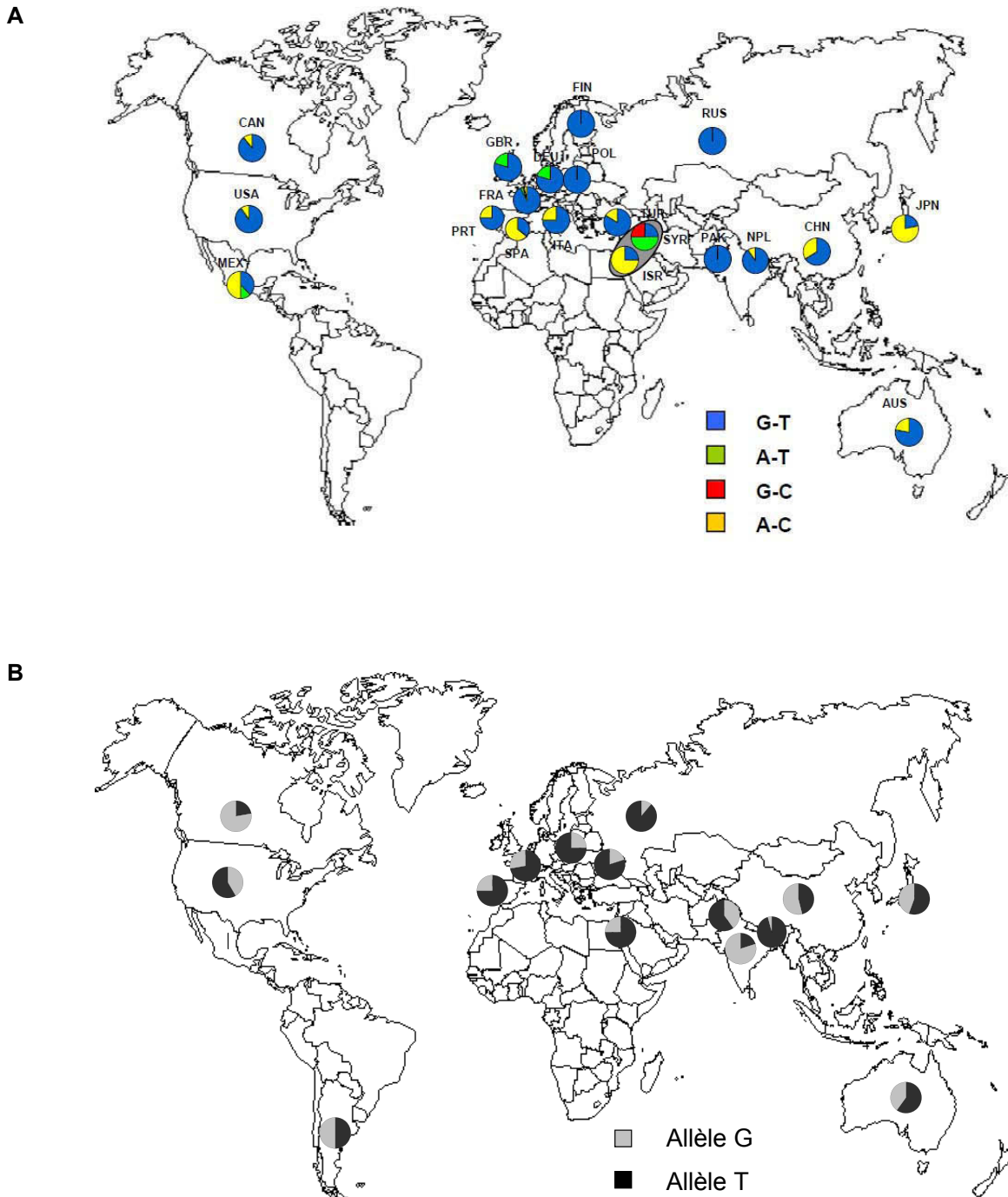


Figure 66 : Distribution géographique mondiale des haplotypes pour TaGW2-A (A) et des allèles identifiés au marqueur pour TaGW2-B (B). Seuls les pays représentés au minimum par quatre variétés ont été pris en compte. (A) La région grisée indique la région du Croissant Fertile, berceau des variétés de blés modernes.

2.3.3. Répartition géographique mondiale des allèles des marqueurs pour *TaGW2-A* et *-B*.

Afin d'estimer la distribution géographique des 4 haplotypes de *TaGW2-A* et des allèles de *TaGW2-B* dans les variétés de la core collection mondiale, seuls les pays de provenance incluant au moins 4 variétés ont été considérés.

Pour *TaGW2-A*, d'une manière générale, l'haplotype G-T a été majoritairement sélectionné à travers le monde (Figure 66, A). Dans la région du Croissant Fertile, berceau d'origine des blés cultivés actuels, il y a coexistence des 4 haplotypes (zone grisée dans la Figure 66, A). En Asie, l'haplotype G-T est le plus fréquent (66-100% des haplotypes présents dans les variétés d'un pays asiatique considéré) excepté au Japon où la situation est inverse avec une fréquence élevée de l'haplotype A-C (77,7%). Dans les accessions européennes, l'haplotype G-T a été majoritairement sélectionné (75-100% des haplotypes présents dans les variétés d'un pays européen considéré) excepté en Espagne où l'haplotype A-C est majoritaire (63,6%).

La répartition géographique mondiale des allèles de *TaGW2-B* est plus homogène et moins contrastée entre les pays de provenance que celle des haplotypes de *TaGW2-A* (Figure 66, B). L'allèle T est l'allèle majoritairement sélectionné excepté au Canada, en Inde et en Chine.

L'ensemble de ces résultats de génétique d'association montre que du polymorphisme dans la séquence de *TaGW2-A* est majoritairement associé à des variations phénotypiques du poids de 1000-grains, de la quantité d'azote dans le grain, de la teneur en protéines du grain et de la dureté du grain, dans la core collection mondiale de blé tendre. Une faible association a également été relevée entre *TaGW2-A* et la longueur de l'épi. L'haplotype G-T de *TaGW2-A* est toujours l'haplotype favorable pour l'ensemble de ces caractères. De plus, cette analyse d'association a également révélé une association importante entre le polymorphisme au marqueur *cfe273* et la longueur de l'épi, le poids de 1000-grains, la quantité d'azote dans le grain et la teneur en protéines du grain ; et une association faible avec la dureté du grain. De manière intéressante *cfe273* est un marqueur dérivé d'EST, donc d'une région codante exprimée d'un gène chez le blé tendre. Une analyse blastx dans NCBI de la séquence EST montre une homologie de séquence avec une sous-unité du complexe translocase de la membrane interne de la mitochondrie (TIM) chez le riz (e-value 6e-55).

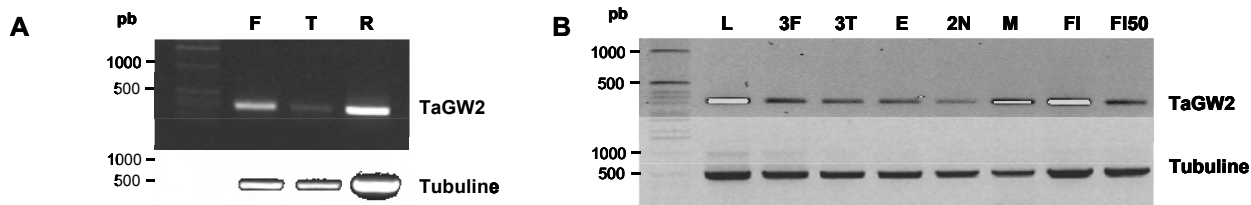


Figure 67 : Accumulation du transcrit TaGW2 dans différents organes (A) et dans des feuilles prélevées à différents stades des développements végétatif et reproductif (B) de *T. aestivum* cv Récital. L analyse a été réalisée par PCR semi-quantitative, à l aide du couple d amorces TaGW2-318. Le gène de la β -tubuline a été utilisé comme gène de référence. Organes en (A): F, T, R, respectivement feuilles, tiges, racines de plantules âgées de 3 semaines. Stades en (B): L, levée; 3F, 3 feuilles; 3T, 3 tiges; E, épiaison; 2N, 2 nœuds; M, méiose; FI, floraison; FI50, 50JAA.

Taux d'expression relative

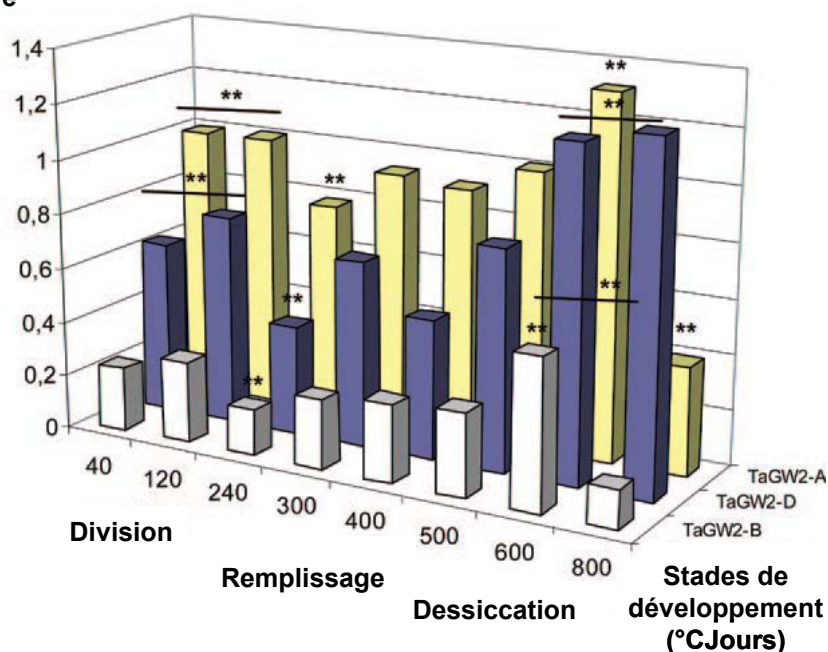


Figure 68 : Taux d'expression relative des transcrits *TaGW2-A*, *-B* et *-D* au cours du développement du grain de blé tendre, cv Récital. La quantification a été réalisée par PCR quantitative sur 8 stades de développement de grains, exprimés en °CJours cumulés. Les valeurs d'expression relative ont été calculées selon la méthode du $2^{-\Delta\Delta Ct}$, en utilisant le gène *Ta2776* comme gène de référence, et les valeurs de Ct au stade 40 °CJours comme calibrateur interne. Toutes les données sont des moyennes de trois réplicats biologiques indépendants. La comparaison de l'accumulation du transcrit entre les stades de développement a été réalisée indépendamment pour chaque copie, en utilisant le test de comparaison multiple de Scheffé. **, différence significative à $p < 0,01$.

3. Analyse de l'expression de *TaGW2*

3.1. Accumulation organe-spécifique des transcrits *TaGW2*

Afin de déterminer si le gène *TaGW2* était exprimé de manière constitutive tout comme son homologue chez le riz, nous avons étudié par PCR semi-quantitative (RT-PCR) l'accumulation des transcrits *TaGW2*, toutes copies confondues, dans différents organes. Le gène *TaGW2* est exprimé dans les racines et les tiges de plantules de blé tendre âgées de 2 semaines, ainsi que dans les feuilles de blé tendre tout au long du cycle de développement de la plante (développements végétatif et reproductif ; Figure 67).

3.2. Profils d'accumulation des transcrits homéologues *TaGW2-A*, *-B* et *-D* au cours du développement du grain

Afin d'analyser précisément le profil d'expression de chacune des copies homéologues de *TaGW2* au cours du développement du grain, nous avons réalisé une analyse par PCR quantitative, sur huit stades de développement de grain de blé tendre variété Récital allant de 40°CJours à 800°CJours, à l'aide de couples d'amorces homéologues-spécifiques (couple d'amorces *TaGW2-A*, *-B* et *-D*). Les transcrits des 3 copies de *TaGW2* sont détectés dès le début du développement du grain (Figure 68). Cependant, les copies *TaGW2* présentent des profils d'expression différents : *TaGW2-A* et *TaGW2-D* ont des profils d'accumulation de transcrits similaires au cours du développement du grain, et *TaGW2-B* diffère des deux précédents. De plus, *TaGW2-B* est peu exprimé dans le grain comparativement aux copies A et D de *TaGW2*. Les transcrits *TaGW2-A* et *TaGW2-D* sont majoritairement exprimés au début de la phase de division du grain (de 40 à 120°CJours) et à la fin du stade de remplissage du grain (600°CJours). Le taux de transcrit *TaGW2-A* chute brutalement à la dessiccation du grain (800°CJours) tandis qu'un taux élevé de *TaGW2-D* est maintenu à ce stade. Les transcrits *TaGW2-B* sont majoritairement exprimés à la fin de la phase de remplissage du grain (600°CJours).

Ce résultat indique tout d'abord que les trois copies homéologues de *TaGW2* sont exprimées tout au long du développement et montre que leurs profils d'expression diffèrent. Les profils d'expression de *TaGW2-A* et *TaGW2-D* suggèrent fortement leur implication au cours des phases de division et la fin du remplissage. De plus, le taux élevé de transcrit *TaGW2-D* suggère que cette copie ait un rôle dans la dessiccation du grain. *TaGW2-B* n'a probablement qu'une implication dans la fin du remplissage en réserve.

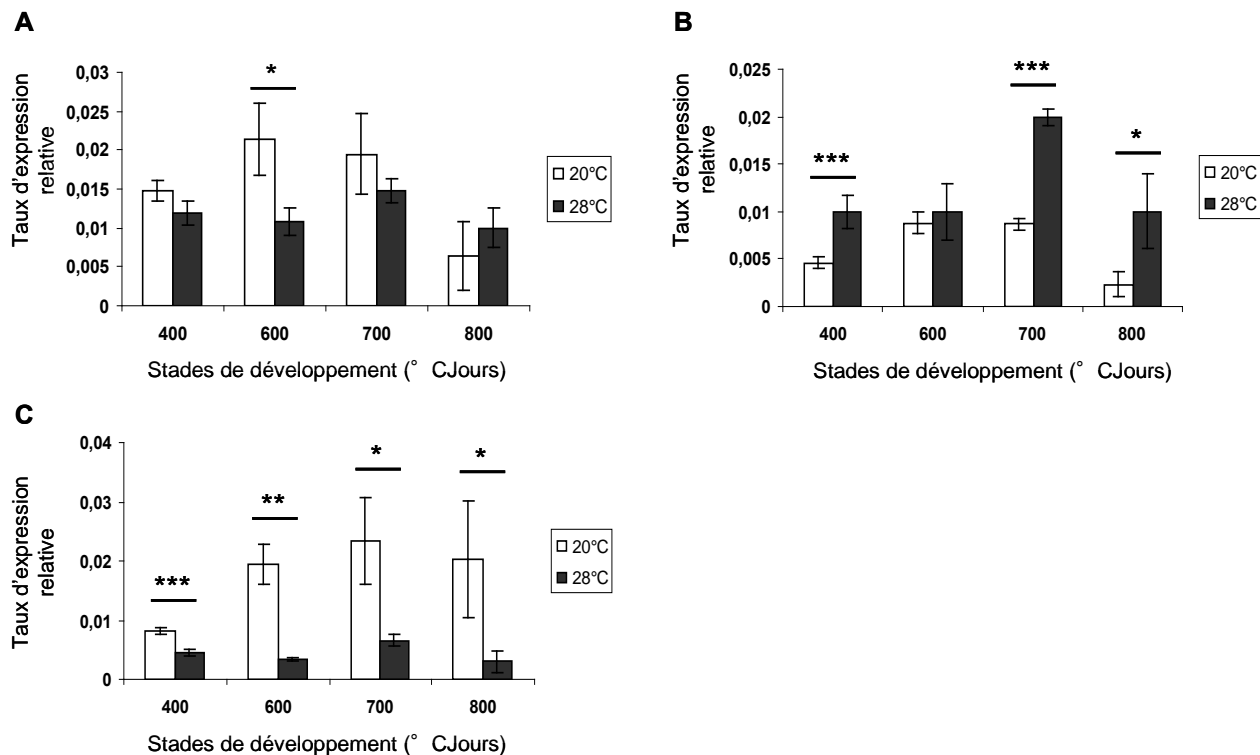


Figure 69 : Taux d'expression relative des transcrits *TaGW2-A* (A), *-B* (B) et *-D* (C) au cours de la phase de remplissage du grain, dans des grains entiers de plantes ayant été placées à 20° C et 28° C. La quantification a été réalisée par PCR quantitative sur 4 stades de développement de grains, exprimés en ° CJours cumulés. Les valeurs d'expression relatives ont été calculées selon la méthode du $2^{-\Delta\Delta Ct}$, en utilisant le gène *Ta2776* comme gène de référence. Toutes les données sont des moyennes de trois réplicats biologiques indépendants. ***, différence significative à $p < 0,001$; **, $p < 0,01$; *, $p < 0,05$ (Test de Student).

3.3. Profils d'accumulation des transcrits homéologues *TaGW2-A*, *-B* et *-D* au cours du remplissage du grain, en conditions de stress thermique.

Tout au long de ce travail de thèse, plusieurs cinétiques de développement de grain ont été développées au laboratoire. Bien que les conditions de température soient contrôlées dans la serre, les cinétiques de développement de grain mises en place à la période estivale souffrent généralement de stress thermique, et sont alors écartées des études. Une cinétique « stressée » a tout de même été conservée : les grains avaient subi une température moyenne journalière de $20\pm 2^{\circ}\text{C}$ jusqu'au stade de développement d'environ 200°CJours ; puis une température moyenne journalière de 28°C a été enregistrée. Des grains entiers aux stades de développement 400, 600, 700 et 800°CJours ont été prélevés, et le profil d'accumulation des transcrits *TaGW2-A*, *-B* et *-D* a été analysé par PCR quantitative (Figure 69). Les copies homéologues de *TaGW2* semblent répondre différemment à un stress thermique appliqué au cours de la phase de remplissage du grain. Le profil d'accumulation du transcrit *TaGW2-A* en conditions « stressée » n'est pas significativement différent de celui en conditions optimales de croissance (20°C), excepté à 600°CJours où une diminution de *TaGW2-A* en conditions « stressée » est observée. De manière contrastée, les copies B et D de *TaGW2* répondent très fortement à un stress thermique appliqué au cours de la phase de remplissage du grain : l'accumulation des transcrits *TaGW2-B* est significativement augmentée, tandis que la quantité relative de transcrits *TaGW2-D* chute fortement.

Bien que cette expérience nécessite d'être réitérée en conditions contrôlées afin de valider ces premiers résultats, cette première analyse semblent montrer que les copies homéologues de *TaGW2* répondent différemment à l'application d'un « stress thermique » au cours de la phase de remplissage du grain.

Au vu de l'ensemble des résultats précédents et des études de Su *et al* (2011) et Yang *et al* (2012), il apparaîtrait intéressant de caractériser en priorité les protéines TaGW2-A et -D. Malheureusement, au moment de ces travaux de thèse, nous n'avons obtenu la séquence ADNc complète que des copies *TaGW2-A* et *-B*, et uniquement une séquence partielle de *TaGW2-D*. Nous nous sommes donc focalisés sur la caractérisation de la protéine TaGW2-A.

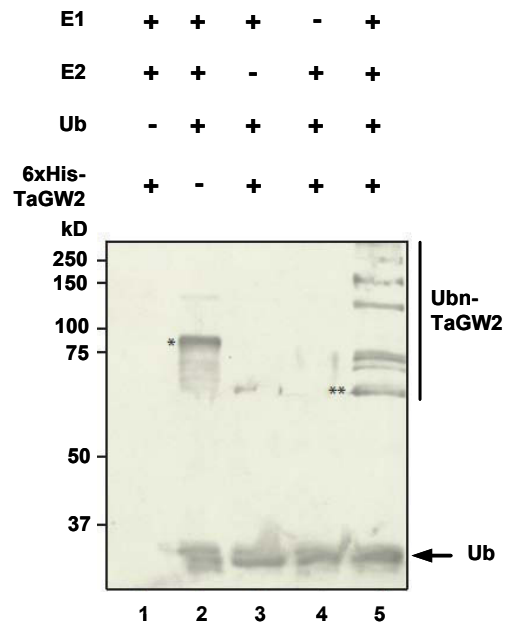


Figure 70 : Activité E3 ubiquitine ligase de la protéine TaGW2-A. La protéine fusion 6xHis-TaGW2-A a été testée pour sa capacité d'auto-ubiquitination en présence ou en absence des enzymes humaines E1(A1S9) et E2 humaine (GST-UbcH5b), et d'ubiquitine (GST-Ub). L'ubiquitine (Ub) et les protéines ubiquitinées (Ubn-TaGW2) ont été détectées par immunodétection avec un anticorps anti-Ub (P4D1, Santa Cruz). *, intermédiaires monoUb-E2 ; **, monoUb-TaGW2-A

4. Caractérisation de la protéine TaGW2-A

4.1. Activité E3 ligase *in vitro* de la protéine TaGW2-A

La réaction d'ubiquitination d'une protéine nécessite *in vivo* l'intervention des 3 enzymes E1, E2 et E3. Dans cette réaction, l'enzyme E3 catalyse le transfert de monomères d'ubiquitine sur la protéine cible. *In vitro*, en absence de substrat protéique, les monomères d'ubiquitine peuvent rester attachés à l'enzyme E3 faute de pouvoir être transférés. On parle de phénomène d'autoubiquitination. Afin de déterminer si la protéine TaGW2-A est une E3 ligase fonctionnelle tout comme GW2 chez le riz, nous avons testé sa capacité d'autoubiquitination. Une protéine fusion 6xHis-TaGW2-A a été produite en bactérie, purifiée en conditions dénaturantes et renaturée. La protéine fusion 6xHis-TaGW2-A a été placée dans des milieux réactionnels dans lequel il manquait un des partenaires nécessaires à la réaction et dans un milieu complet. Une immunodétection anti-ubiquitine des réactions montre qu'il y a apparition de protéines ubiquitinées de haut poids moléculaire, uniquement en présence d'un milieu réactionnel complet (Figure 70, piste 5). Les poids moléculaires des bandes protéiques détectées dans la piste 5 correspondent aux poids moléculaires de la protéine fusion 6xHis-TaGW2-A avec un monomère d'ubiquitine (60,2 kDa) ou plus. La protéine fusion 6xHis-TaGW2-A est capable de s'auto-ubiquitiner en présence d'Ub et des enzymes E1 et E2 dans le milieu.

Cependant, une bande protéique ubiquitinée apparaît dans la piste 2 (Figure 70). Il est fort probable que cette bande protéique corresponde à des intermédiaires E2-Ub puisque le poids moléculaire de cette bande détectée correspond au poids moléculaire de l'enzyme E2 avec un monomère d'ubiquitine (77 kDa). En effet, en absence d'E3 dans le milieu réactionnel, l'ubiquitine ne peut pas être transférée à l'enzyme E3 et resterait liée sur l'enzyme E2. Ainsi, l'absence de la E1 ou de la E2 abolit l'autoubiquitination de 6xHis-TaGW2-A.

Ces résultats montrent que la protéine TaGW2-A présente une activité E3 ligase *in vitro*, tout comme son homologue GW2 chez le riz.

4.2. Localisation subcellulaire de la protéine TaGW2-A

Afin de rechercher la fonction de la protéine TaGW2-A au sein de la cellule, nous avons d'abord analysé sa localisation subcellulaire. La protéine TaGW2-A a été fusionnée en partie C-terminale et dans le cadre de lecture de la protéine fluorescente YFP (Yellow

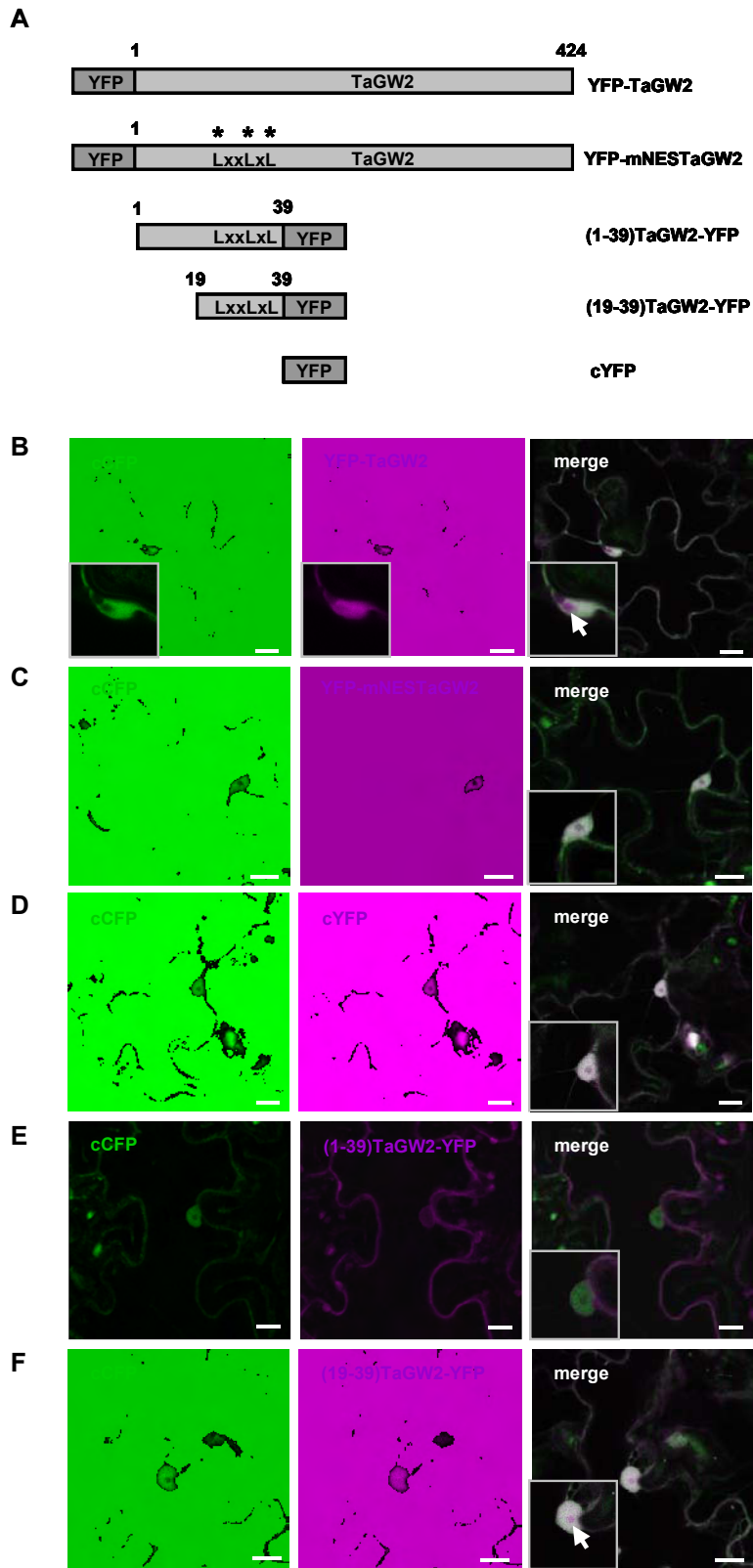


Figure 71 : Localisation subcellulaire de la protéine TaGW2-A et identification d'un signal NES fonctionnel.
(A) Représentation schématique des protéines fusion. **(B-F)** Images de microscopie confocale de cellules d'épiderme inférieur de feuilles de *Nicotiana tabacum* co-exprimant la protéine contrôle cCFP (à gauche) et les protéines fusion YFP (au milieu). Les images superposées des deux canaux (merge) sont présentées à droite. (suite page suivante)

Fluorescent Protein ; Figure 71, A) ; et sa distribution subcellulaire a été examinée par microscopie confocale. La protéine YFP-TaGW2 a été co-exprimée de manière transitoire avec la protéine contrôle cytosolique CFP (cCFP, cytosolic Cyan Fluorescent Protein) dans des cellules d'épiderme inférieur de feuilles de tabac, par agroinfiltration. Les protéines YFP-TaGW2 et cCFP s'accumulent dans le noyau et le cytoplasme (Figure 71, B, merge). YFP-TaGW2 est également détectée dans le nucléole (Figure 71, B, fléché dans l'encart de l'image merge). La localisation nucléaire de cCFP résulte de sa diffusion passive à travers les pores nucléaires du fait de sa petite taille (22,3 kDa). La diffusion passive des protéines étant généralement limitée à des protéines possédant un poids moléculaire inférieur à 40 kDa (Wagstaff et Jans, 2009), le poids moléculaire de la protéine fusion YFP-TaGW2 (69,6 kDa) exclut sa diffusion passive dans le noyau. La protéine TaGW2-A présente donc réellement une localisation nucléaire.

4.3. Identification de domaines responsables de la distribution subcellulaire de la protéine TaGW2-A

Le fait que YFP-TaGW2 soit à la fois détectée dans le noyau et le cytoplasme nous a alors amené à déterminer les motifs protéiques qui pouvaient gouverner sa localisation intracellulaire. A l'heure actuelle, nous ignorons encore comment TaGW2-A entre dans le noyau, puisqu'aucun signal de localisation nucléaire (NLS) canonique n'a pu être détecté dans la séquence primaire de TaGW2-A. Cependant, une région riche en leucine, le motif ³⁴LxxLxL³⁹, a été identifiée en partie N-terminale de la séquence primaire de TaGW2-A, à l'aide du programme NetNES (<http://www.cbs.dtu.dk/services/NetNES>). Les motifs LxxLxL ont été décrits dans la littérature comme agissant en qualité de signal d'export nucléaire (NES, Nuclear Export Signal ; pour exemples Fukuda *et al*, 1996 ; Taagepera *et al*, 1998 ; Toyoshima *et al*, 1998 ; Wada *et al*, 1998). Afin de déterminer si ce domaine putatif NES est fonctionnel, nous avons généré des mutants de délétions et de substitutions de TaGW2-A fusionnés à YFP (Figure 71, A). Dans le mutant de substitution YFP-mNESTaGW2, les trois leucines ont été remplacées par des alanines (³⁴AxxAxA³⁹). De manière cohérente avec la destruction d'un domaine NES fonctionnel, YFP-mNESTaGW2 s'accumule exclusivement dans le noyau (Figure 71, C). Ce résultat indique que le motif ³⁴LxxLxL³⁹ est nécessaire pour l'export nucléaire de YFP-TaGW2, et que la destruction de ce motif conduit à une rétention nucléaire de la protéine chimère.

Afin de déterminer si le motif ³⁴LxxLxL³⁹ est suffisant pour l'export nucléaire de TaGW2-A, nous avons ensuite testé la région protéique des acides aminés 1 à 39 de TaGW2-

(Suite de la légende de la Figure 71) : Une vue détaillée du noyau est donnée dans les encarts. YFP-TaGW2 (**B**) co-localise avec cCFP dans le cytoplasme et le noyau. YFP-TaGW2 s'accumule également dans le nucléole (pointé par une flèche dans l'encart de l'image merge). YFP-mNESTaGW2 (**C**) est retenue dans le nucléoplasme et est exclue du nucléole. cYFP (**D**) s'accumule dans le cytoplasme et le nucléoplasme. (**E**) Les 39 premiers acides aminés de TaGW2-A sont capables de rediriger la protéine cYFP du noyau vers le cytoplasme. (**F**) L'addition des acides aminés 19 à 39 de TaGW2-A en partie N-terminale de YFP résulte en l'accumulation de (19-39)TaGW2-YFP dans le nucléole. Son accumulation dans le cytoplasme et le nucléoplasme reste inchangée par rapport à cYFP. Agrandissement x20. Barre, 10 μ m.

A, contenant le signal NES putatif ³⁴LxxLxL³⁹, pour l'export nucléaire de la protéine cYFP ((1-39)TaGW2-YFP). La protéine cYFP seule, tout comme la protéine cCFP, s'accumule dans le cytoplasme et diffuse passivement dans le noyau (Figure 71, D). La protéine (1-39)TaGW2-YFP est majoritairement détectée dans le cytoplasme (Figure 71, E), bien que sa taille (26,8 kDa) autoriserait sa diffusion passive dans le noyau. La région des acides aminés 1 à 39 de TaGW2-A est donc suffisante pour exporter cYFP du noyau. Cependant, les acides aminés 19 à 39 de TaGW2-A ((19-39)TaGW2-YFP) ne sont pas suffisants pour exporter cYFP du noyau vers le cytoplasme, puisque la protéine fusion (19-39)TaGW2-YFP est détectée dans le noyau et le cytoplasme (Figure 71, F). Ce résultat suggère que la conformation tridimensionnelle de la région protéique comportant le motif LxxLxL est probablement importante pour que ce motif NES soit accessible à la machinerie cellulaire d'export nucléaire.

De manière intéressante, la protéine fusion YFP-mNESTaGW2 est exclue du nucléole (Figure 71, C, merge). Ce résultat suggère que le motif LxxLxL pourrait aussi agir comme un signal de rétention nucléolaire. En effet, la fusion des acides aminés 19 à 39 de TaGW2-A, contenant le motif LxxLxL, à YFP résulte clairement en l'accumulation de la protéine fusion (19-39)TaGW2-YFP dans le nucléole (Figure 71, F, fléché dans l'encart de l'image merge). Cependant, les 39 premiers acides aminés de TaGW2-A fusionnés à YFP n'entraînent pas une accumulation nucléolaire de (1-39)TaGW2-YFP (Figure 71, E). Il est important de rappeler que (1-39)TaGW2-YFP est activement exportée du noyau et s'accumule majoritairement dans le cytoplasme. Nous pouvons émettre l'hypothèse que (1-39)TaGW2-YFP ne peut pas s'accumuler dans le nucléole du fait d'une trop faible proportion de la protéine fusion (1-39)TaGW2-YFP restante dans le nucléoplasme pour être adressée au nucléole.

Dans la littérature, des régions riches en résidus glycine et arginine (domaines GAR, Glycine-Arginine Rich) et des signaux de localisation nucléolaire (NoLS), riches en acides aminés basiques, ont été montré dans la rétention nucléolaire de protéines (Schmidt-Zachmann et Nigg, 1993 ; Heine *et al*, 1999 ; Emmott et Hiscox, 2009). Les domaines GAR sont généralement de grandes bandes de répétitions du motif GRR (plusieurs dizaines de répétitions). Il n'existe pas de séquence canonique pour les NoLS : le signal de localisation nucléolaire de la protéine kinase NIK (NF-κB-inducing Kinase) par exemple est un motif RKKRKKK (Birbach *et al*, 2004), tandis que celui de l'isoforme étendue de la protéine C/EBPα est un motif RRXR (Müller *et al*, 2010). Dans les premiers acides aminés de la séquence primaire de TaGW2-A, nous avons identifié un motif glycine-arginine, le motif ⁴RxGGRR⁹. Ce motif est un mélange des deux types de signaux de rétention nucléolaire

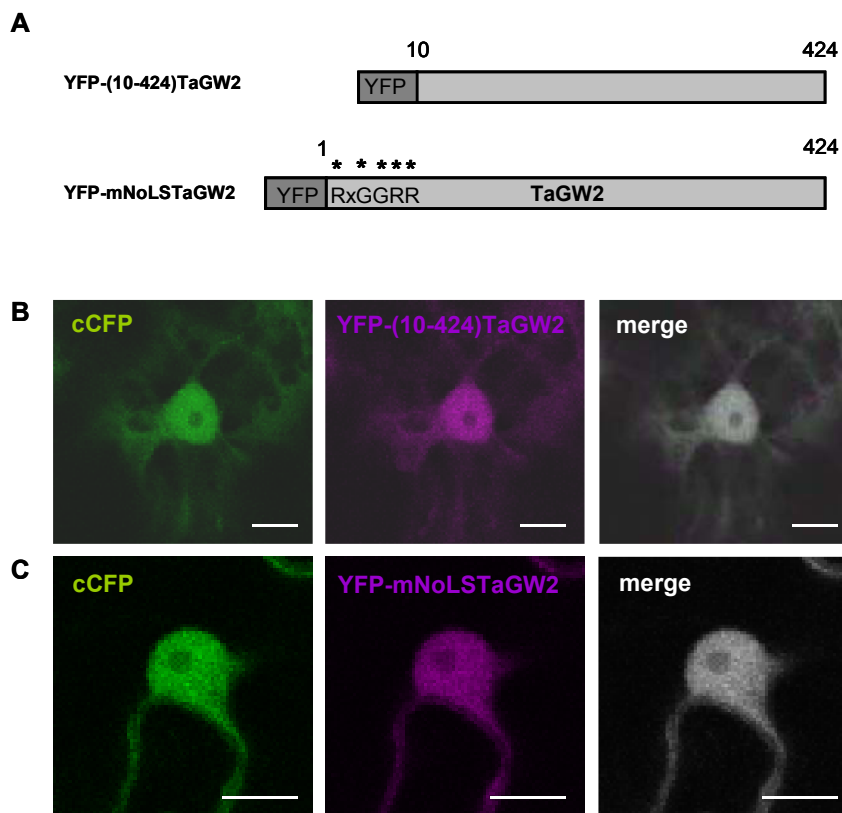


Figure 72 : Identification d'un signal NoLS fonctionnel dans TaGW2-A. (A) Représentation schématique des protéines fusion YFP. (B) YFP-(10-424)TaGW2 s accumule dans le nucléoplasme et le cytoplasme. La délétion des acides 1 à 9 de TaGW2 entraîne une exclusion nucléolaire de la protéine YFP-(10-424)TaGW2. (C) YFP-mNoLS s accumule dans le nucléoplasme et le cytoplasme, et est exclue du nucléole. Barre, 10 μ m.

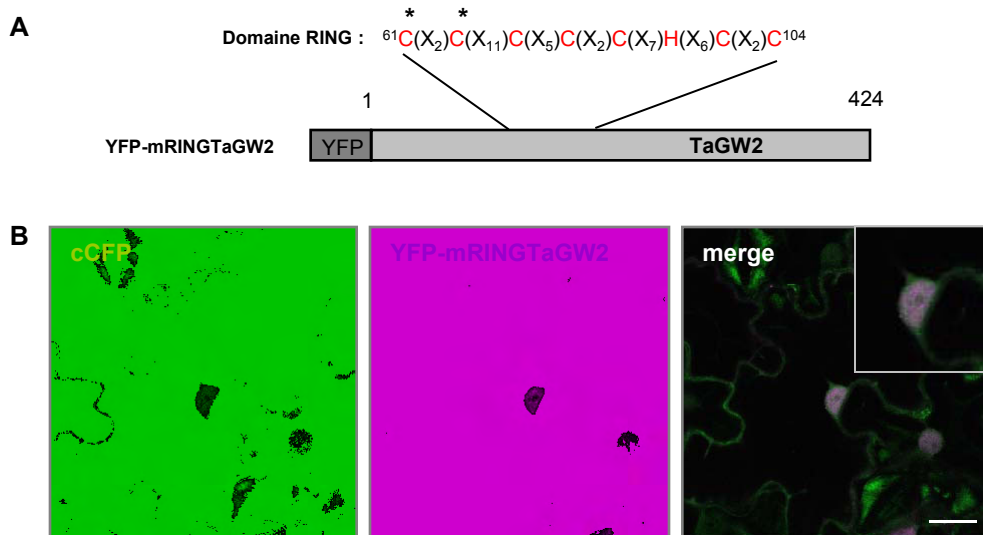


Figure 73 : La perte de fonction E3 ligase de la protéine TaGW2-A modifiée sa localisation subcellulaire. (A) Représentation schématique du mutant de substitution pour le domaine RING fusionné en C-terminale de YFP. *, acides aminés mutés. (B) YFP-mRINGTaGW2 est retenue dans le nucléoplasme. Barre, 20 μ m.

précédemment cités : il possède les résidus types d'un domaine GAR mais sa longueur ne permet pas de le classer dans cette catégorie de domaines et le rapproche plus du signal NoLS des protéines NIK et C/EBP α étendue. Afin de déterminer si ce motif est un NoLS fonctionnel, nous avons généré un mutant de délétion et un mutant de substitution NoLS, de TaGW2-A, fusionnés à YFP (Figure 72, A). Le mutant YFP-(10-424)TaGW2 est délété pour la région en acides aminés 1 à 9 de TaGW2-A, contenant le motif ⁴RxGGRR⁹. Dans le mutant de substitution YFP-mNoLSTaGW2, tous les résidus arginine et glycine du motif ont été remplacés par des alanines (⁴AxAAAA⁹). YFP-(10-424)TaGW2 s'accumule dans le nucléoplasme et le cytoplasme, mais n'est pas détecté dans le nucléole (Figure 72, B). De plus, de manière cohérente avec la destruction d'un domaine NoLS fonctionnel, YFP-mNoLSTaGW2 ne s'accumule pas dans le nucléole, mais seulement dans le nucléoplasme et le cytoplasme (Figure 72, C). Ces résultats indiquent que le motif ⁴RxGGRR⁹ est nécessaire pour la localisation nucléolaire de YFP-TaGW2, et que la destruction de ce motif conduit à une exclusion du nucléole de la protéine chimère.

L'ensemble de ces résultats semblent montrer donc que les motifs ³⁴LxxLxL³⁹ et ⁴RxGGRR⁹ sont des signaux de localisation nucléolaire de TaGW2-A, fonctionnant de manière indépendante.

Enfin, afin de déterminer si une perte de l'activité E3 ligase de la protéine TaGW2-A peut modifier la localisation subcellulaire de la protéine, nous avons généré un mutant de substitution de TaGW2-A muté pour le domaine RING, fusionné à YFP (YFP-mRINGTaGW2). Les cystéines en position 61 et 64 de la séquence primaire de TaGW2-A ont été remplacées par des arginines (Figure 73, A). La protéine fusion YFP-mRINGTaGW2 s'accumule exclusivement dans le nucléoplasme (Figure 73, B). Ce résultat montre que les cystéines en position 61 et 64 du domaine RING de TaGW2-A sont importantes pour les localisations cytoplasmique et nucléolaire de TaGW2-A. Ceci suggère que la fonction E3 ligase serait reliée à la localisation subcellulaire de TaGW2-A.

Conclusions et Discussions du Chapitre I

En conclusion de cette première partie de nos travaux, nous avons montré que *TaGW2* est exprimé par trois copies, portées par chacun des génomes homéologues A, B et D du blé tendre. Ces copies exprimées de *TaGW2* sont très proches puisqu'elles partagent 98% d'identité de séquence en nucléotides. Les copies génomiques de *TaGW2* sont situées sur le chromosome 6 ; et les copies *TaGW2-A* et *TaGW2-B* ont été cartographiées plus précisément sur le bras court des chromosomes 6A et 6B, proche du centromère. Du polymorphisme non-codant (promoteur et intron 7) dans la copie de *TaGW2-A* a été retrouvé associé au poids de 1000-grains, à la dureté et à la quantité d'azote et de protéines dans le grain, dans une core-collection mondiale de blé tendre. Dans notre étude, l'haplotype G-T aux marqueurs TaGW2_A_R(-593) et TaGW2_A_Y1061 respectivement est toujours l'allèle favorable pour chacun des caractères précédemment cités, et a majoritairement été sélectionné dans les variétés de blé tendre à travers le monde.

Le gène *TaGW2* est exprimé de manière constitutive. Les 3 copies homéologues *TaGW2* sont exprimées tout au long du développement du grain de blé tendre. Cependant, elles présentent des profils d'accumulation de transcrit différents en conditions optimales de développement du grain : les copies *TaGW2-A* et *-D* sont les plus exprimées dans le grain et présentent un profil d'accumulation similaire, suggérant leur implication dans la phase de division et la transition remplissage/dessiccation du grain. La copie *TaGW2-B* est beaucoup moins exprimée au cours du développement du grain et présente un maximum d'accumulation du transcrit uniquement à la transition remplissage/dessiccation du grain. Les 3 copies homéologues semblent se comporter également différemment dans l'expression de leur transcrit au cours de la phase de remplissage, suite à l'application d'un « stress thermique » au cours de cette phase de développement. L'expression de *TaGW2-A* est faiblement affectée par ce traitement tandis que les accumulations des transcrits *TaGW2-B* et *-D* sont respectivement augmentées et diminuées dans ces conditions.

Au cours de ces travaux, nous avons également montré que la protéine TaGW2-A est une E3 ubiquitine ligase fonctionnelle *in vitro*, et qu'elle présente une localisation nucléaire, nucléolaire et cytoplasmique. En partie N-terminale de la protéine TaGW2-A, un motif LxxLxL est responsable à la fois de l'export de la protéine du noyau vers le cytoplasme et de son adressage au nucléole. De plus, un motif RxGGRR en partie N-terminale de TaGW2-A est également responsable de sa rétention nucléolaire. Pour finir, il semble qu'une perte de sa fonction E3 ligase modifie sa localisation subcellulaire.

Tableau 10: QTLs pour le rendement en grain et ses composantes reportés sur le chromosome 6AS. Les QTLs pour lesquels le marqueur TaGW2_A_Y1061 se retrouve dans leur intervalle de confiance sont surlignés en jaune. La correspondance des marqueurs avec la carte consensus nous a également permis d'exclure TaGW2-A de certains QTLs [a]. Pour les autres QTLs, ne connaissant que la position d'un seul marqueur ou n'ayant pas de correspondance entre les marqueurs et la carte consensus, nous ne pouvons conclure quant à la présence de TaGW2-A sous ces QTLs [b].

Référence	Composante	Population(s)	Nbre d'environnements	Intervalle de marqueurs ou marqueurs les plus proches	Position sur carte consensus (si correspondance possible)	
Kuchel 2007	Rendement en grain	Trident x Molineux	1/18	Xwmc163	39.8	[b]
	Poids de 1000-grain	Trident x Molineux	2/3	Xwmc163	39.8	
Sun 2009	Largeur du grain	131 individus RIL de 'Chuan 35050' x 'Shannong 483'	2/4	swes123 - swes332	nd	[b]
	Poids de 1000-grain	131 individus RIL de 'Chuan 35050' x 'Shannong 483'	4/4	gwm617 - swes332	nd	
Gegas 2010	Largeur du grain	Avalon x Cadenza	1/2	wPt-9690 - barc0107	30.38 - 38	[a]
	Poids de 1000-grain	Savannah x Rialto	1/1	wPt-2573 - gwm1005	30.19 - ?	
	Largeur du grain	Savannah x Rialto	1/1	wPt-2573 - gwm1005	30.19 - ?	
	Poids de 1000-grain	Malacca x Charger	1/1	psp3152 - wmc179	35.41 - 54	
	Taille du grain*	Avalon x Cadenza	nd	gdm36 - barc113	sous-intervalle entre les marqueurs wmc182 et wPt8833: 29 - 33.07	
	Taille du grain*	Savannah x Rialto	nd	gwm334 - gwm169	sous-intervalle entre les marqueurs wmc182 et wPt8833: 29 - 33.07	
Sun 2010	Poids spécifique	132 individus RIL de 'Ning7840' x 'Clark'	2/7	wmc807 - gwm132	41 - 48	[a]
	Poids du grain	132 individus RIL de 'Ning7840' x 'Clark'	5/7	wmc398 - gwm132	37 - 48	[a]
	Poids du grain	132 individus RIL de 'Ning7840' x 'Clark'	4/7	barc1055 - wmc807	? - 41	[b]
	Diamètre du grain	132 individus RIL de 'Ning7840' x 'Clark'	4/7	ctcg.gac1 - wmc398	? - 37	[b]
	Diamètre du grain	132 individus RIL de 'Ning7840' x 'Clark'	4/7	wmc398 - gwm132	37 - 48	[a]
Heidari 2011	Rendement en grain	Fukuho-kumogi' x 'Oligoculm'	1/2	p02m22 - p02m3	nd	[b]
	Rendement en grain	Fukuho-kumogi' x 'Oligoculm'	1/2	p10m31 - p10m10	nd	
Wang 2012	Poids de 1000-grain	262 accessions d'une mini core collection de blés chinois	1/5	gwm494	35.8	[b]
	Poids de 1000-grain	262 accessions d'une mini core collection de blés chinois	3/5	gwm356	45.12	

Les résultats de l'analyse de l'expression de *TaGW2* ne seront pas discutés dans cette partie mais seront mis en lien avec les résultats de validation fonctionnelle, dans le chapitre II.

TaGW2-A est situé dans une région génomique du 6AS influençant des composantes du rendement et de qualité du grain.

La copie A de *TaGW2* a été cartographiée en partie proximale du bras court du chromosome 6A, dans la population Ct x CS, au sein d'une région génomique où plusieurs QTLs pour le rendement ou des composantes du rendement ont précédemment été rapportés (Marza *et al*, 2006 ; Kuchel *et al*, 2007 ; Sun *et al*, 2009 ; Gegas *et al*, 2010 ; Sun *et al*, 2010 ; Heidari *et al*, 2011 ; Wang *et al*, 2012). Nous avons recensé dans la littérature l'identification de 6 QTLs pour le poids de 1000-grains, 3 QTLs de rendement en grains, 3 QTLs de largeur du grain, 2 QTLs pour le poids du grain, 2 QTLs pour le diamètre du grain et 1 QTL pour le poids spécifique en grain, sur le chromosome 6AS du blé tendre, dans différentes populations (Tableau 10). Certains marqueurs bornant ou couvrant les intervalles de confiance de ces QTLs sont communs à la carte consensus dont nous disposons. Connaissant la position du marqueur TaGW2_A_Y1061 sur la carte consensus (32,7), nous avons alors pu déterminer si *TaGW2-A* se situait dans l'intervalle de confiance de ces QTLs. *TaGW2-A* est situé dans l'intervalle de marqueurs d'1 QTL de poids de 1000-grains, de 2 QTLs de largeur du grain et de 2 QTLs pour la taille du grain, tous reportés dans l'analyse de Gegas et ses collaborateurs (2010) (Tableau 10, surlignés en jaune).

Nos résultats de génétique d'association suggèrent que la région chromosomique incluant les marqueurs cfe273 et TaGW2_A_R(-593) définit un QTL pour le poids de 1000-grains. Une analyse d'association entre les 3 marqueurs et des données phénotypiques obtenues pour un sous-ensemble de 193 variétés de la CC372, dans 2 environnements supplémentaires, a révélé une faible association ($p=0,04$) du marqueur TaGW2_A_R(-593) avec le poids de 1000-grains (Tableau 11). On peut noter que dans cette analyse, il ne semble pas y avoir d'effet marqueur x environnement. Toutefois, dans notre étude de liaison sur la population Ct x CS, nous n'avons détecté aucune co-localisation de *TaGW2-A* avec un QTL de poids de 1000-grains. Il est intéressant de rappeler que plusieurs QTLs pour le poids de 1000-grains ont été reportés sur le chromosome 6AS mais pratiquement aucuns n'ont été retrouvés dans tous les environnements testés lors des études considérées ; et aucune région méta-QTL pour des composantes du rendement, y compris de poids de 1000-grains, n'a été identifiée jusqu'à présent sur le chromosome 6A (Zhang *et al*, 2010). L'ensemble de ces

Tableau 11 : Analyse d'association multilocale entre le polymorphisme pour *TaGW2-A* et le marqueur *cfe273*, et la variation phénotypique pour le poids de 1000-grains et les quantité d'azote et de protéines dans le grain, dans 193 individus de la core collection mondiale de blé tendre. L analyse phénotypique de ce sous-groupe de variétés de la CC372 a été réalisée à Clermont-Ferrand (2006), et au Moulon (2007) avec deux niveaux d intrants. Le modèle GLM utilisé comprend l effet marqueur, l effet environnement et l interaction marqueur x environnement. Les valeurs mentionnées sont des moyennes des valeurs observées dans l ensemble des variétés pour un allèle. Les résultats significatifs sont surlignés en jaune.

Marqueur	cfe273				TaGW2_A_R(-593)				TaGW2_A_Y1061			
Allele	303	336	p-value marqueur	p-value marqueur x env	G	A	p-value marqueur	p-value marqueur x env	T	C	p-value marqueur	p-value marqueur x env
Poids de 1000-grains (g)	41,190	43,366	0,381	0,875	43,647	40,105	0,041	0,746	43,503	41,827	0,498	0,372
Quantité d'azote par grain (mg d'azote par grain)	0,828	0,941	0,015	0,773	0,950	0,851	0,015	0,213	0,944	0,871	0,115	0,166
Teneur en protéines du grain (% de la masse sèche du grain)	11,195	12,169	0,007	0,765	12,200	11,854	0,222	0,006	12,200	11,646	0,090	0,004

données et résultats suggère donc que le QTL poids de 1000-grains identifié sur le 6AS n'est pas stable et est probablement soumis à l'influence de l'environnement.

D'autre part, l'analyse multilocale sur le sous-ensemble de 193 variétés confirme les régions QTLs pour les quantités d'azote et de protéines du grain mises en évidence par l'analyse d'association unilocale sur l'ensemble des 372 variétés : les marqueurs *cfe273* et *TaGW2_A_R(-593)* ont été retrouvés significativement associés à la teneur en azote du grain, et le marqueur *cfe273* significativement associé à la teneur en protéines du grain ; sur les 2 environnements (Tableau 11). Cette région sur le 6AS incluant les marqueurs *cfe273*, *TaGW2_A_R(-593)* voire *TaGW2_A_Y1061* semble donc définir également une région QTL pour la quantité en azote et protéines du grain. Ces résultats sont cohérents avec l'étude de Charmet et ses collaborateurs (2005) qui reporte la présence de QTLs pour la masse sèche du grain, la vitesse d'accumulation d'azote et la quantité de protéines dans le grain en partie proximale du 6AS. Ces études d'association proposeraient donc d'autres caractères potentiellement impactés par *TaGW2-A*, à savoir la quantité d'azote et de protéines dans le grain. Cependant, étant donné l'étendue du déséquilibre de liaison dans des espèces cultivées comme le blé tendre, nous ne pouvons pas exclure que *TaGW2-A* ne soit pas réellement impliqué dans le contrôle de ces paramètres de remplissage du grain et que les deux marqueurs dans *TaGW2-A* soient simplement liés au(x) polymorphisme(s) causal situé(s) dans un autre gène. Ceci d'autant plus que le marqueur *cfe273* apparaît être un meilleur candidat dans nos analyses d'association uni- et multilocales pour expliquer les quantités d'azote et de protéines dans le grain. Toutefois, Song et ses collaborateurs (2007) ont montré que *GW2* chez le riz impactait le poids du grain en contrôlant également, en plus de la largeur du grain, l'accumulation en réserves au cours de la phase de remplissage. La validation fonctionnelle de *TaGW2* est par conséquent essentielle afin de connaître précisément le rôle de *TaGW2*.

En conclusion de cette partie, l'ensemble de ces résultats suggère que cette région du 6AS contient probablement plusieurs gènes, dont *TaGW2-A* et le gène dont le marqueur *cfe273* est issu, et contrôle plusieurs caractères agronomiques. Notre étude confirme l'association entre du polymorphisme pour *TaGW2-A* et le poids de 1000-grains mise en évidence par Su et ses collaborateurs ; mais enrichit également le set de caractères potentiellement impactés par *TaGW2-A*, à savoir la teneur en azote et protéines du grain de blé.

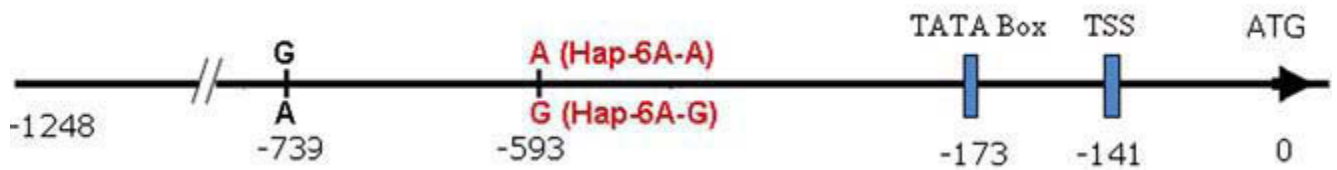


Figure 74: Positions et allèles des deux marqueurs développés dans le promoteur de *TaGW2-A*, dans l'étude de Su et ses collaborateurs (2011). TSS, site d'initiation de la transcription; ATG, codon d'initiation de la traduction.

Comparaison de l'étude de Su et al (2011) et de notre étude d'association.

Dans leur étude, Su et ses collaborateurs ont reporté une association entre le marqueur au promoteur de *TaGW2-A*, TaGW2_A_R(-593), et le poids de 1000-grains, dans une core collection de blé chinois. Nos résultats de génétique d'association sont cohérents avec l'étude de Su *et al* puisque le marqueur TaGW2_A_R(-593) a également été retrouvé associé au poids de 1000-grains, dans la core collection mondiale de blé tendre. Cependant, l'autre marqueur que nous avons développé pour la copie A de *TaGW2*, TaGW2_A_Y1061, n'a pas été retrouvé associé à ce caractère. Ces résultats apparemment contradictoires pour deux marqueurs d'un même gène peuvent être expliqués par le déséquilibre de liaison non-absolu observé entre ces deux marqueurs. En effet, le déséquilibre de liaison entre les marqueurs TaGW2_A_R(-593) et TaGW2_A_Y1061, observé dans la population mondiale de blé tendre, était de 0,7. La présence de recombinants pour ces deux marqueurs suggère donc un(des) crossing-over(s) au sein du gène *TaGW2-A*. Chez le blé tendre, Saintenac et ses collaborateurs (2009, 2011) ont montré une corrélation entre la présence de gènes et la distribution de crossing-overs sur une portion du chromosome 3B. Un(des) crossing-over(s) au sein de *TaGW2-A* pourrait être une explication plausible au déséquilibre de liaison non-absolu entre ces 2 marqueurs pour *TaGW2-A*.

Cependant, bien que notre analyse confirme l'association précédemment mise en évidence par Su (2011) du marqueur TaGW2_A_R(-593) au poids de 1000-grains, l'allèle favorable pour le caractère était l'allèle G dans notre étude, contre l'allèle A dans l'étude précédente. Deux hypothèses peuvent être formulées quant à cette divergence entre les études. La première est que les études d'association ont été menées sur des populations différentes : Su et ses collaborateurs ont mené leur analyse sur une population de blé chinois, regroupant des populations de landraces et de variétés modernes ; notre analyse a été réalisée sur une population mondiale de blé tendre. De manière intéressante, Su et ses collaborateurs mentionnent dans la partie discussion de leur article que bien qu'ils aient identifié l'allèle A au marqueur TaGW2_A_R(-593) comme étant l'allèle favorable, « néanmoins, des variétés possédant l'allèle G à ce marqueur présentaient aussi des grains avec une largeur et un poids élevés ». La deuxième hypothèse est que nous ayons fait une inversion dans les jeux de données de génotypage. Notre étude étant conduite après l'étude de Su (2011) et amenant à des résultats contradictoires des résultats de cette dernière, nous avons minutieusement vérifié à plusieurs reprises notre stratégie de génotypage, i.e. les génotypes révélés par le couple d'amorces que nous avons développé pour le génotypage de TaGW2_A_R(-593), et nos

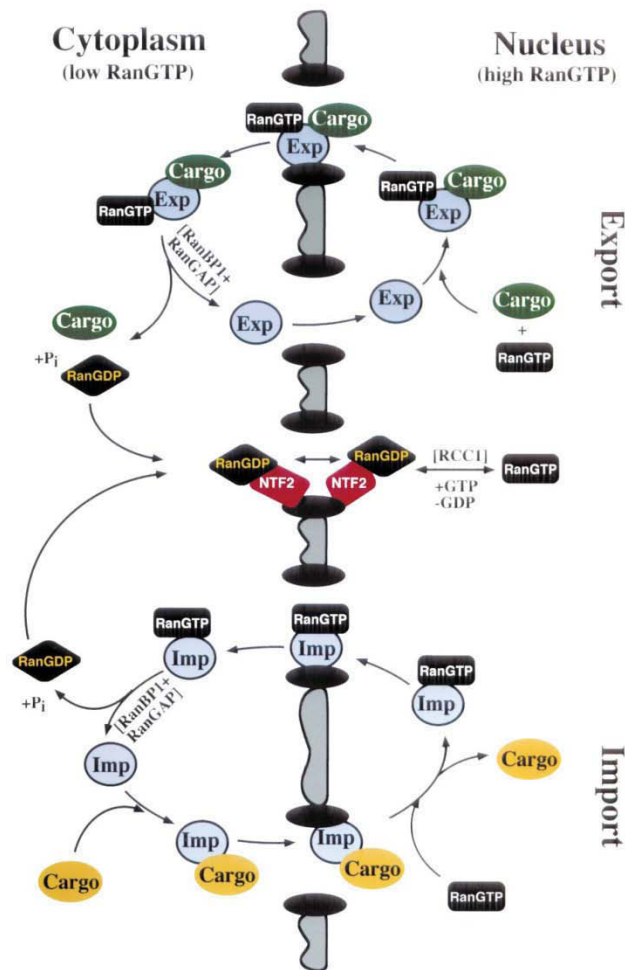


Figure 75: Le transport nucléo-cytoplasmique actif de protéines à travers les pores de la membrane nucléaire (d'après Görlich et Kutay, 1999). Imp, importines; Exp, exportines; Cargo, protéine possédant un signal d'import ou d'export nucléaire.

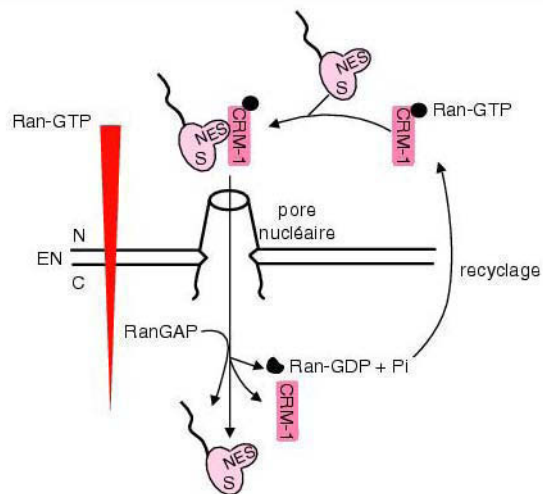


Figure 76: L'export nucléaire de protéines dépendant du récepteur CRM1 (d'après Sitterlin, 2005). NES, signal d'export nucléaire; S, substrat protéique; N, noyau; C, cytoplasme; EN, enveloppe nucléaire.

données de génotypage pour ce marqueur. Il s'avère que dans notre étude d'association l'allèle G est l'allèle favorable notamment pour le caractère poids de 1000-grains. Il peut être important de rappeler que Su et ses collaborateurs ont mentionné dans leur étude un marqueur SNP en position -739 en parfait déséquilibre de liaison avec le marqueur TaGW2_A_R(-593), dans le promoteur de TaGW2-A (Figure 74). Tout comme pour TaGW2_A_R(-593), les allèles possibles au marqueur TaGW2_A_R(-739) sont G et A. Les deux marqueurs étant en déséquilibre de liaison parfait d'après les auteurs, on peut donc en déduire que l'allèle favorable au marqueur TaGW2_A_R(-739) pour le poids de 1000-grains est G. Il est possible que les auteurs aient malencontreusement inversé les jeux de données et relevé l'allèle favorable au marqueur en position -739 et non en position -593.

TaGW2-A est une protéine nucléo-cytoplasmique

Nous avons montré que la protéine TaGW2-A présente une localisation nucléaire, nucléolaire et cytoplasmique. Dans cette étude, deux domaines responsables du trafic intracellulaire de TaGW2-A ont été identifiés : un domaine NES de type LxxLxL et un domaine NoLS.

Le transport de macromolécules entre le noyau et le cytoplasme est un processus finement régulé. Leur passage à travers la membrane nucléaire a lieu à travers les complexes du pore nucléaire (NPC). Bien que le diamètre des NPC autorise la diffusion passive de petites protéines (<40 kDa), la plupart des protéines présentant une fonction nucléaire apparaissent activement transportée dans et hors du noyau. Le transport nucléo-cytoplasmique actif est un processus signal-dépendant assuré par une famille de récepteurs/transporteurs appelés récepteurs karyopherin ou importines/exportines (Ström et Weis, 2001 ; Görlich et Kutay, 1999 ; Figure 75). Plusieurs voies d'export nucléaire ont été identifiées jusqu'à présent, dont la plus connue et la plus étudiée est la voie d'export nucléaire dépendante d'un signal NES riche en leucine. Ce signal a été montré dans de nombreux organismes comme étant reconnu par le récepteur karyopherin CRM1 (Fornerod *et al*, 1997 ; Fukuda *et al*, 1997 ; Neville *et al*, 1997 ; Ossareh-Nazari *et al*, 1997 ; Stade *et al*, 1997 ; Haasen *et al*, 1999 ; Figure 76). La voie d'export dépendante de CRM1 est une voie très conservée à travers les organismes eucaryotes (Haasen *et al*, 1999) et requiert l'action coopérative de la protéine RanGTP. Cette voie peut être efficacement et spécifiquement inhibée par le fongicide Leptomycin B (Fornerod *et al*, 1997). Le motif LxxLxL identifié en partie N-terminale de TaGW2-A correspond aux séquences canoniques de motifs riches en leucine reconnus par la voie d'export nucléaire dépendante de CRM1 (La Cour *et al*, 2004). Afin de vérifier que

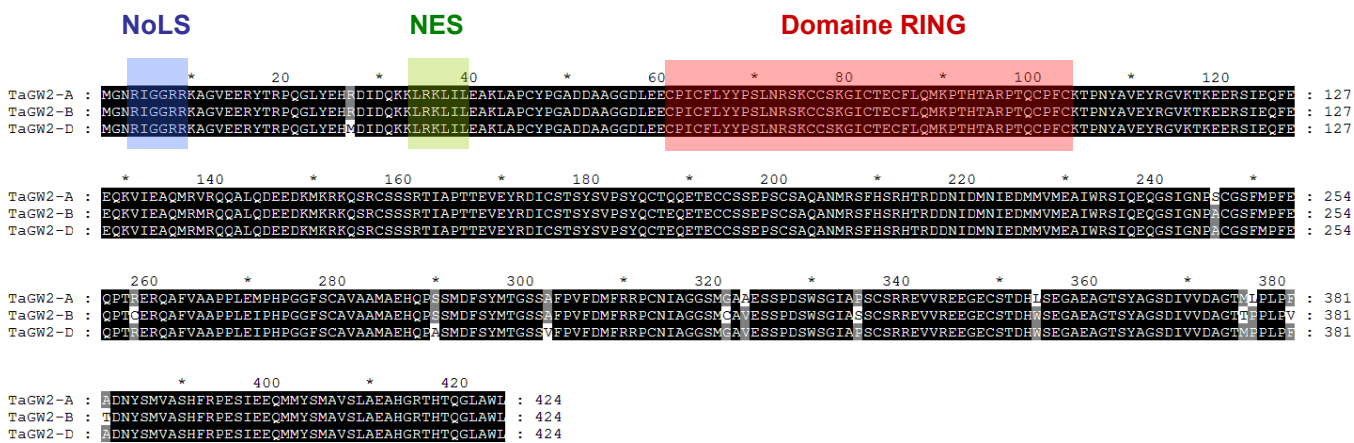
TaGW2-A est exportée vers le cytoplasme par la voie CRM1, nous avons initié des traitements à la leptomycine B (LMB) sur des feuilles de tabac agroinfiltrées exprimant des constructions d'intérêt de manière transitoire. Nous avons tout d'abord tenté de définir les conditions d'expérimentation sur des feuilles de tabac exprimant de manière transitoire une protéine fusion contrôle, (GFP)-NLS-chalcone synthase (CHS)-NES (gracieusement fournie par le Pr Merkle, Université de Bielefeld), qui est activement exportée vers le cytoplasme par la voie CRM1. A ce jour, nous n'avons pas encore réussi à déterminer la concentration et le mode d'administration de LMB sur un système d'expression transitoire de protéine fusion. Nous avons donc initié récemment la transformation stable pour YFP-TaGW2 de cellules de tabac BY-2 afin de pouvoir effectuer ces traitements LMB, dans le but de voir si TaGW2 est exportée vers le cytoplasme par une voie dépendante de CRM1.

Le nucléole est une sous-structure au sein du noyau. Il n'est pas physiquement délimité par une membrane mais peut être considéré comme un sous-compartiment du noyau à part entière, du fait qu'il peut être identifié morphologiquement par microscopie électronique ou encore isolé biochimiquement, et qu'il contient un sous-ensemble défini de macromolécules lui étant propres. Le nucléole n'étant pas entouré par une membrane, la localisation nucléolaire de protéines contraste avec l'import nucléaire qui nécessite un transport actif des protéines à travers les NPC. Il est suggéré dans la littérature que les protéines nucléolaires s'accumulent dans ce compartiment par affinité pour des sites de liaisons au sein de celui-ci (Misteli, 2001) ; et par conséquent les domaines de localisation nucléolaire doivent être considérés plus comme des signaux de rétention plutôt que des signaux d'import actif. Lors de notre étude, nous avons caractérisés fonctionnellement deux signaux de rétention au nucléole au sein de la protéine TaGW2 : un motif riche en arginine et glycine NoLS classique (AA4-9) et de manière intéressante, le motif LxxLxL assurant également la fonction d'export nucléaire. Ces deux motifs semblent indépendants et ne semblent pas être deux parties d'un même domaine puisque 1) l'ajout des acides aminés 19 à 39 de TaGW2 en N-terminale de la protéine YFPc (protéine chimère (1-39)TaGW2-YFP, donc sans le NoLS des AA 4 à 9) est suffisant pour retenir YFPc au sein du nucléole, 2) la seule mutation du NoLS (protéine chimère YFP-mNoLSTaGW2) permet d'inhiber la rétention nucléolaire de TaGW2. Jusqu'à présent, nous n'avons pas véritablement recensé dans la littérature de tel cas de double fonction (export nucléaire et rétention nucléolaire) pour un domaine NES de type LxxLxL. Néanmoins, deux hypothèses peuvent être formulées quant à la nature de ce signal de rétention nucléolaire atypique. La première hypothèse est que le

véritable signal de rétention nucléolaire dans cette région de TaGW2 est un motif plus large que le motif LxxLxL que nous avons mis en évidence au cours de notre étude. En effet, dans cette partie de la séquence primaire de TaGW2, le motif ³⁴LxxLxL³⁹ est inclus dans un motif d'acides aminés basiques plus large ³²KKLRKLxL³⁹, qui ressemble à un motif NoLS. Le signal NoLS dans cette région de la protéine TaGW2 et le NES serait donc imbriqué l'un dans l'autre. Les mutations de substitutions (leucine en alanine) dans le mutant mNESTaGW2 auraient alors également modifié la conformation et détruit le NoLS dans cette région. La deuxième hypothèse repose sur la reconnaissance de ce domaine NES par CRM1. Très récemment, chez les mammifères, des sous-structures dans le nucléole, riches en protéine CRM1 ont été identifiées et appelées CNoBs (CRM1 nucleolar bodies ; Ernoult-Lange *et al*, 2009). Les auteurs ont notamment montré que la protéine nucléo-cytoplasmique CPEB1 (Cytoplasmic polyadenylation element-binding protein 1) pouvait s'accumuler dans le nucléole, au niveau de ces sous-structures, grâce à une interaction avec CRM1 via le(s) domaine(s) NES de la protéine CPEB1. L'export nucléaire de la protéine CPEB1 a également été montré assuré par la voie dépendante de CRM1 via la reconnaissance de ces domaines NES. Toutefois, les structures CNoBS découvertes chez les mammifères par l'équipe de Ernoult-Lange (2009) sont des structures très ponctuelles dans le nucléole ; et l'accumulation nucléolaire de la protéine chimère (19-39)TaGW2-YFP dans notre étude semble occuper tout le nucléole. Les expériences de traitement LMB sur des cellules BY-2 transformées stablement pour YFP-TaGW2 permettront peut-être de trancher entre ces deux hypothèses.

La taille de la protéine chimère YFP-TaGW2-A (69,6 kDa) ne permettant pas sa diffusion passive au travers des NPC, la protéine doit donc être importée dans le noyau par un transport actif. L'import nucléaire classique de protéines implique la reconnaissance de signaux NLS (Nuclear localisation signal) dans la protéine considérée par des récepteurs de la famille des importines. Les NLS classiques sont des bandes de résidus lysine et arginine et peuvent être monopartite ou bipartite (Lange *et al*, 2007). La séquence consensus couramment avancée est K(K/R)x(K/R), mais par exemple la séquence NLS de la protéine virale SV40 large T est ¹²⁶PKKKRRV¹³² ou encore celle de la protéine nucléoplasmine est ¹⁵⁵KK(X₁₀)KKKK¹⁷⁰. Au sein de la séquence primaire de TaGW2-A, seuls deux motifs pouvant agir en tant que NLS classiques ont été identifiés : le motif KxKRK des acides aminés 149 à 153, et le motif KKxRK des acides aminés 32 à 36. Toutefois, comme nous l'avons mentionné dans la partie précédente, le motif ³²KKxRK³⁶ est imbriqué dans le domaine NES fonctionnel (³²KKLRKLxL³⁹) ; et les mutations de substitutions dans le mutant

A



B

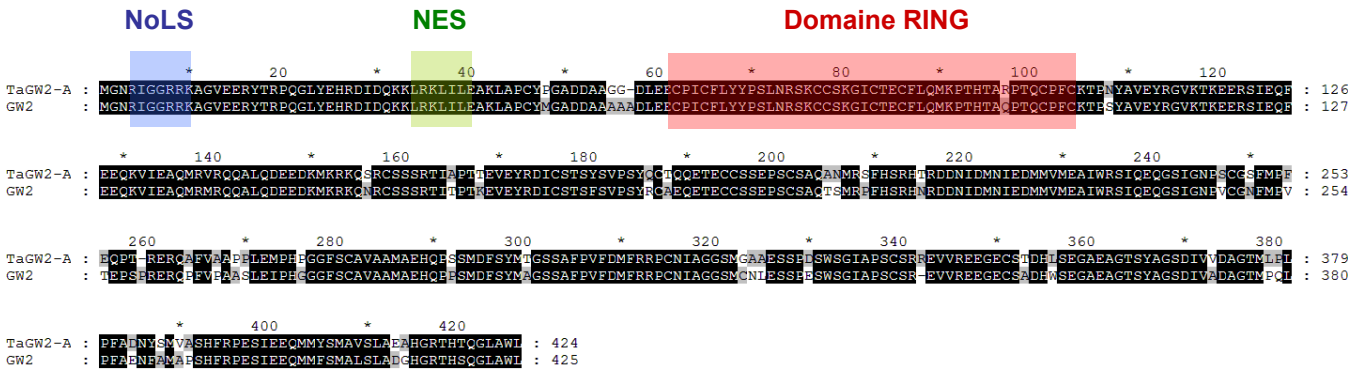


Figure 77 : Alignements des séquences primaires des protéines homéologues TaGW2-A, -B et -D (A), et de TaGW2-A chez le blé tendre et GW2 chez le riz (B). Les acides aminés identiques entre les séquences sont indiquées en noir. Le domaine RING et les domaines NES et NoLS identifiés dans TaGW2-A au cours de cette étude sont figurés respectivement en rouge, vert et bleu.

mNESTaGW2 ont entraîné une exclusion nucléolaire et cytoplasmique mais une accumulation nucléoplasmique de la protéine chimère, et son exclusion. Il est donc fort probable que ce motif (³²KKxRK³⁶) ne soit pas le signal NLS d'adressage de TaGW2-A au noyau. Les constructions de fusion YFP de mutants de substitutions pour ces motifs sont en cours.

Cependant, il est également possible que TaGW2 soit importée au noyau indépendamment des importines. En effet, de plus en plus d'études reportent désormais l'import nucléaire de protéines par des transporteurs non conventionnels ou encore un passage à travers l'enveloppe nucléaire par interaction avec les protéines du NPC (revue de Wangstaff et Jans, 2009)

Il est fort probable que les protéines homéologues TaGW2-B et TaGW2-D présentent également une localisation nucléaire, nucléolaire et cytoplasmique, tout comme TaGW2-A. En effet, les signaux NoLS (⁴RxGGRR⁹) et NES (³⁴LxxLxL³⁹) caractérisés fonctionnellement au cours de cette étude sont strictement conservés entre les séquences primaires des homéologues (Figure 77, A).

De manière contrastée, Song et ses collaborateurs (2007) ont montré que la protéine GW2 présente une localisation cytoplasmique. Cependant, les signaux NoLS et NES identifiés dans TaGW2-A sont également strictement conservés dans la séquence primaire de GW2 (Figure 77, B). Il peut être intéressant de souligner qu'à ce jour, nous n'avons pas encore identifié le(s) domaine(s) nécessaires à l'import de TaGW2-A dans le noyau. L'identification de ce(s) domaine(s) permettra peut être d'expliquer cette différence de localisation subcellulaire des deux homologues, GW2 et TaGW2-A.

TaGW2-A est une E3 ubiquitine ligase

Notre étude a confirmé que la protéine TaGW2-A possédait une activité E3 ligase, tout comme son homologue GW2 chez le riz. De plus, nous avons montré que la mutation du domaine RING de TaGW2-A (YFP-mRINGTaGW2) entraînait une exclusion nucléolaire et cytoplasmique de la protéine chimère, et sa rétention au sein du nucléoplasme. Ceci laisse à penser que TaGW2-A ubiquitine sa(ses) cible(s) protéique(s) au sein du cytoplasme et du nucléole. L'ubiquitination de protéines au sein du cytoplasme et du noyau a été reportée de longue date. Au cours de la dernière décennie, quelques exemples d'ubiquitination de protéines nucléolaires, engendrant soit leur dégradation soit leur stabilisation, ont été reportés (Chen *et al*, 2002 ; Itahana *et al*, 2003 ; Sato *et al*, 2004). Des études ont également montré

une accumulation de protéasomes et de protéines cibles potentielles, ubiquitinées, dans le nucléole, suite à des traitements par des inhibiteurs du protéasome (Mattson *et al*, 2001 ; Arabi *et al*, 2003). Le nucléole est le centre de synthèse des ribosomes, et l'ubiquitination des protéines ribosomiques est fortement reliée au processus de biogenèse des ribosomes (Spence *et al*, 2000 ; Andersen *et al*, 2005 ; Lam *et al*, 2007). En outre, il a été montré qu'une inhibition de l'activité protéasomale pouvait engendrer des défauts importants dans la production d'ARN ribosomiques (ARNr) matures, activité majeure du nucléole (Stravrera *et al*, 2006). Le nucléole est donc un lieu tout aussi important d'ubiquitination des protéines. Toutefois, une rétention nucléolaire de E3 ligases a également été montrée comme mécanisme d'action de rétention de l'activité E3 de ces protéines : par exemple, Ang et ses collaborateurs (1998) ont montré que, chez les animaux, la rétention nucléolaire de la E3 ligase MDM2 suite à un stress hyperprolifératif autorise son substrat protéique, la protéine p53, à agir dans le noyau. Nous n'excluons donc pas que TaGW2-A présente une localisation nucléolaire du fait de sa possible rétention pour le blocage de son activité E3. Cependant, la perte de fonction E3 (mutant RING) entraînant une exclusion nucléolaire de TaGW2-A, il est fort probable que TaGW2-A agisse dans le nucléole.

La fonction première et la plus connue du nucléole est la biogenèse des ribosomes matures. Les ARN ribosomiques (ARNr) y sont transcrits puis assemblés aux protéines ribosomales pour former les sous-unités ribosomales. Des études récentes suggèrent désormais des fonctions supplémentaires du nucléole (Pendle *et al*, 2005 ; revues de Boisvert *et al*, 2007 ; Emmott et Hiscox, 2009 ; Shaw et Brown, 2012). Des protéines à localisation nucléolaire, avec aucune fonction connue dans la biogenèse des ribosomes, ont notamment été montrées dans la régulation du cycle cellulaire et dans les réponses à des stress chez les animaux (San-Segundo et Roeder, 1999 ; Yishida *et al*, 2002 ; Geil *et al*, 2008 ; Suzuki *et al*, 2012).

Chapitre II :
Recherche de la fonction de *TaGW2* dans le
grain de blé tendre en développement par
validation fonctionnelle

L'objectif principal de ce travail de thèse était de valider la fonction du gène *TaGW2* au cours du développement du grain de blé tendre. Chez cette espèce, l'analyse fonctionnelle de gènes souffre de nombreuses limites, causée par la nature hexaploïde de son génome. En effet, beaucoup de gènes, comme c'est le cas pour *TaGW2*, sont portés et exprimés par les trois génomes homéologues. De plus, les copies homéologues de ces gènes sont souvent redondantes fonctionnellement. Des collections de mutants de TILLING sont désormais disponibles chez le blé tendre. Cependant, l'invalidation d'une seule copie du gène peut souvent ne pas être suffisante, l'effet de cette simple invalidation pouvant être masquée par la redondance fonctionnelle des autres copies non-invalidées ; et l'obtention de mutants nécessite souvent l'invalidation simultanée des trois copies du gène. L'obtention de mutants perte-de-fonction chez le blé tendre reste par conséquent plus difficile et plus longue que chez d'autres espèces.

Une alternative est de déduire la fonction d'un gène en utilisant des mutants chez des espèces phylogénétiquement proches du blé tendre, comme le riz ou brachypodium par exemple. Cependant, les phénomènes ancestraux de polyploïdisation chez le blé tendre ont parfois conduit à des différenciations fonctionnelles entre le blé tendre et le riz. De plus, la fréquence importante d'évènements de duplication et de délétion dans les génomes du blé a modifié la colinéarité à certains endroits du génome entre le riz et le blé tendre. De manière identique, des études de génomique comparative ont montré de nombreuses altérations de la colinéarité entre les génomes du blé tendre et de brachypodium. Ces changements peuvent donc altérer les relations fonctionnelles entre des orthologues chez le blé tendre et ces 2 espèces ; et rendre difficile la prédiction de la fonction d'un gène chez le blé tendre à partir de la validation fonctionnelle chez l'une de ces 2 espèces. *In fine*, une validation fonctionnelle du gène chez le blé tendre semble donc toujours nécessaire.

L'extinction de gènes par ARN interférence permet de supprimer simultanément l'expression de ce type de gène à chacun des loci (la technique d'ARN interférence est détaillée dans la partie Matériel et méthodes, p.73-76). Afin d'étudier *in planta* la fonction du gène *TaGW2* chez le blé tendre, nous avons donc tout d'abord générés des plantes transgéniques sous-exprimant les trois copies du gène *TaGW2*, par ARN interférence.

De plus, d'après nos travaux de génétique d'association et ceux de Su et ses collaborateurs (2011), la copie A de *TaGW2* semble impacter le poids du grain de blé. Nous avons donc également initié la génération de plantes transgéniques surexprimant la copie A de *TaGW2*.

Figure 78: Extrait des séquences génomiques de *T. aestivum* ciblées par la cassette RNAi *TaGW2*. Ces séquences ont été extraites de la base de données de l'URGI. Pour chaque séquence identifiée (ou « job »), sont représentés en bleu les régions codantes pour un gène, et en rouge et en gras la position de l'identité de séquence continue de 20 nt au minimum avec la séquence partielle de *TaGW2-A*. Pour le job #7, le résultat de l'analyse blastx est donné.

Au cours de ce travail de thèse, seules les lignées transgéniques sous-exprimant *TaGW2* ont pu être phénotypées et analysées.

1. Production des lignées RNAi *TaGW2*.

1.1. Analyse des séquences génomiques ciblées par la cassette RNAi *TaGW2*.

Afin d'éteindre l'expression du gène *TaGW2* par ARN interférence, une cassette permettant la production d'ARN double-brin de *TaGW2* a été construite. Un fragment de 1242 pb de la séquence codante de *TaGW2-A* (de la base 34 à 1275 du CDS de *TaGW2-A*) a été inséré en orientation sens et antisens dans le vecteur pSTARGATE, en aval d'un promoteur fort constitutif.

Afin de vérifier que la séquence utilisée pour éteindre les gènes *TaGW2* n'engendre pas l'extinction d'autres gènes, nous avons procédé à une recherche de séquences présentant une identité de séquence continue de 20 nt au minimum avec la séquence insérée dans la cassette RNAi *TaGW2*, dans les bases de données de séquences génomiques de blé tendre disponibles.

Tout d'abord, une analyse blastn dans la base de données de l'URGI (<http://urgi.versailles.inra.fr>) a permis l'identification de 8 séquences présentant une identité de séquence continue de 20 nt avec la séquence partielle de *TaGW2-A*. Parmi ces 8 séquences, 4 séquences correspondent aux copies *TaGW2-A*, *TaGW2-B* et *TaGW2-D* sur les chromosomes 6AS, 6BS et 6DS respectivement. Les 4 autres séquences ne correspondent à aucune des séquences des 3 copies de *TaGW2* (Figure 78). Le job #5 correspond à une séquence sur le chromosome 7AL présentant une identité de séquence continue de 22 nt avec la séquence partielle de *TaGW2-A* ; le job #6, une séquence sur le chromosome 3AL avec 20 nt d'identité continue avec la séquence partielle de *TaGW2-A* ; et les job#7 et #8 des séquences sur le 2AS avec 22 nt et 20 nt d'identité continue avec la séquence partielle de *TaGW2-A*, respectivement. Afin de déterminer si ces « bandes » d'identité de séquence continue sont localisées dans des régions codantes de gènes, une analyse blastx dans NCBI (<http://blast.ncbi.nlm.nih.gov>) a été réalisée sur les séquences de chacun de ces 4 jobs. Lorsque qu'il y en a, les régions exoniques de gènes ont été délimitées dans la séquence du job et sont illustrées en bleu dans la Figure 78. Une identité de séquence continue de 20 nt au minimum entre la séquence partielle de *TaGW2-A* et une séquence exonique d'un autre gène que les copies *TaGW2* a été trouvée uniquement dans le job#7. L'analyse blastx dans NCBI

Figure 78 (suite)

#job_7 >gnl|IWGSC_2AS|IWGSC_chr2AS

Longueur de la séquence = 7655 pb

Identité de séquence avec la séquence partielle de TaGW2-A :

identité = 22/22 (100%), Gaps = 0/22 (0%)

Query 894 CCGCCGACCGTGCAACATTGCT 915

|||||

Sbjct 5994 CCGCCGACCGTGCAACATTGCT 5973

TCCCTCGAGGTGACAAATTTGTTATGTGTTGCGGCTGATGTTTCCTTGCACTAACACGACGCGGAATATGAAGCGTTGCTCCATGGTCTTCGAATGGCT
AAATAGATGAGCCTGAGCTGGGTCCGGTGCTTTGGCGACTCAGATTTGGTAGCTCAACAGGTGTGAGCAAGTGGGATTCCGAAGATCCTCTCATGGCGGCT
TATCACCCTGAAGTTGATGCTGTTGCAGGCGATTTTAAAGGTTATCAGGTGGAGCACATTGACCGTAGAAAAGACAAAGCGGCTGATGCTTTAAGCCGATTG
GGATCTCAGCGTAAGCCGGTGCCACCTAACACCTTTTGGATGTTTGCATAACCCGCTCTGTTAAGTTACCTACAGAAGAAGATCTAGCCGTTCTGACCCA
GAGCGCAATTGGTGGCAGCTCTTCACGTATCCAGATTGGACAGTACCGTTCTTGGCTTACATGACCCGGAGCGAGCTGCCTGTGGATGAAACCCCTGGCC
CGACAGATAACCCGGCGGTCTAAGTCAATGACGATTTTGGACGAGAACTATACCATCGTAGTGTCTCTGGAGCATTTACAGCGGTGTGTTTCCCCTGAAGAA
GGCCAAGAAATACTTCGTGAGATCCATGAAGGTGATTGTGGTATCACGCCGGGTCAAATCTCTGGTGGCCAAAGCTTTTCGTGATGGTTTTATTGGTTG
ATGGCTAACGCTGATGCAGAAGATCTAGTTAGTACATGTGATGGATGTCAAAAATTCGCACGACGCGCACATGTGTCGGCTCAAGAATTGCGGATGATTCCA
ATCACTTGGCCGTTTGGGCTCTGGGGCTTGACATGGTTGGACCTTTCAAAAGGTCTAAGGATAAAAAGACACATCTTTAGTGGCAGTCGACAAATTCACA
AAGTGGGTTGAGGCAGAACCAAGTGAGTAAGTGTGATGCGGCCACAGTGGTTCAAGTTTCAAGAAAAGGTGATCTCCGTTTTGGTTTTCCACATAGCATTAGC
ACTGAAAATGGTACAAATCTATCCCAAGGGGCTATGGAGGAGTTCTGTGACGCGAGCATATCCGGCTTGATGTCTCTTCTGTAGCTCATCTCAATCCAAT
GGTCAGGCTGAGAGAGCAAATCAGGAAATATTGAAAGGTCTCAAGCCCGGCTTATGGTTCCCTTATAGTGAACGCCGGGTTGTTGGGTAGAGGAGTTACCC
TCGGTGCTGTGGAGTATCAACACCACCCCTAACAGATCTACAGGTTATGCACCTTCTTCATGGTTTATGGAGCGGAAGCAGTGTTCCTAGTGACATCCGT
CATGACTCGCCCGTGTGGCGGCTTATGTTGAAGCTAACAAATGAACAAGCCAGACAGGATGCACCTTGACTTGTTAGATGAGGAAAGAGACTTAGCAGCGGCT
CGATTAGCAATTTATCAACAGGACCTACGCCGTTATCACAGCCGCCGGGTTAAGTCCAGGAATTTTCAAGAAGGTGACCTGGTGCTCTGGCTCATCGAGGAT
CTATCAGATGCACACAAGCTATCCCCACATTGGGAAGGACCCTTTGTGGTGAGCAAGAACTTAAACAATGGGTGATATTACCTCATTGACGTTTCGAGAGCAC
AAAGACTCATGTAAATCGGAAGAGAGAGACCGCCGACCGTGCAACATTGCTCAACTTCGTCCTACTACACCTGAGCTATCGGCTCTCATCATGTACATACG
TTGACGTTATATATATATATGATAAGTAATAAGCAGGACCATCGGTCTCTTTTCTTTTCAGAAAAGATTACACATCTTTATTATTTTTACTACTATTTTCAA
ATGACTATTAAGGAGCTGATCATATTTGAATCAAGTCTAACCTTTTTTGGTCCGGCTTATGATCGTATTTCGAA...ACTGCCGTATGTAAGGAAAGACTTGTA
AGAAAGGCATGTAAAGAAATCACCAAGCCAAATCAATTTATGAGCCGGCAGGACTCTGTAGGCCACAGGC

Résultat de l'analyse blastx dans NCBI:

Accession	Description	Max score	Total score	Query coverage	E value	Max ident
ABG65979.1	retrotransposon protein, putative, Ty3-gypsy subclass [Oryza sativa Japonica Group]	480	480	22%	2,00E-143	46%

#job_8 >gnl|IWGSC_2AS|IWGSC_chr2AS

Longueur de la séquence = 257 pb

Identité de séquence avec la séquence partielle de TaGW2-A :

identité = 27/29 (94%), Gaps = 1/29 (3%)

Query 1094 TGCTACCGTTGCCTTTTGCTGACAATTAC 1122

|||||

Sbjct 177 TGCTACCGTTGCCTTTTGCT-ACCATTAC 204

GAATTGGAGTAGTTCTGTGTACCCCTAGGTTTTGATTGTTACATGATGAACCGCATCCGGCATAATTCTCTATCACCAGATCCATTACCTACGAGCATTCCA
TATATTGTTCTTCGCTTATTTACTTTTCCATTGCTATTGTTATCATACTACAAAACACCAAAATATTACTTTTGCTACCGTTGCCTTTTGCTACCATAC
CACTACTATCATATTACTTTGCTACTAAACACTTTGCTGCAGATATTAAGTTT

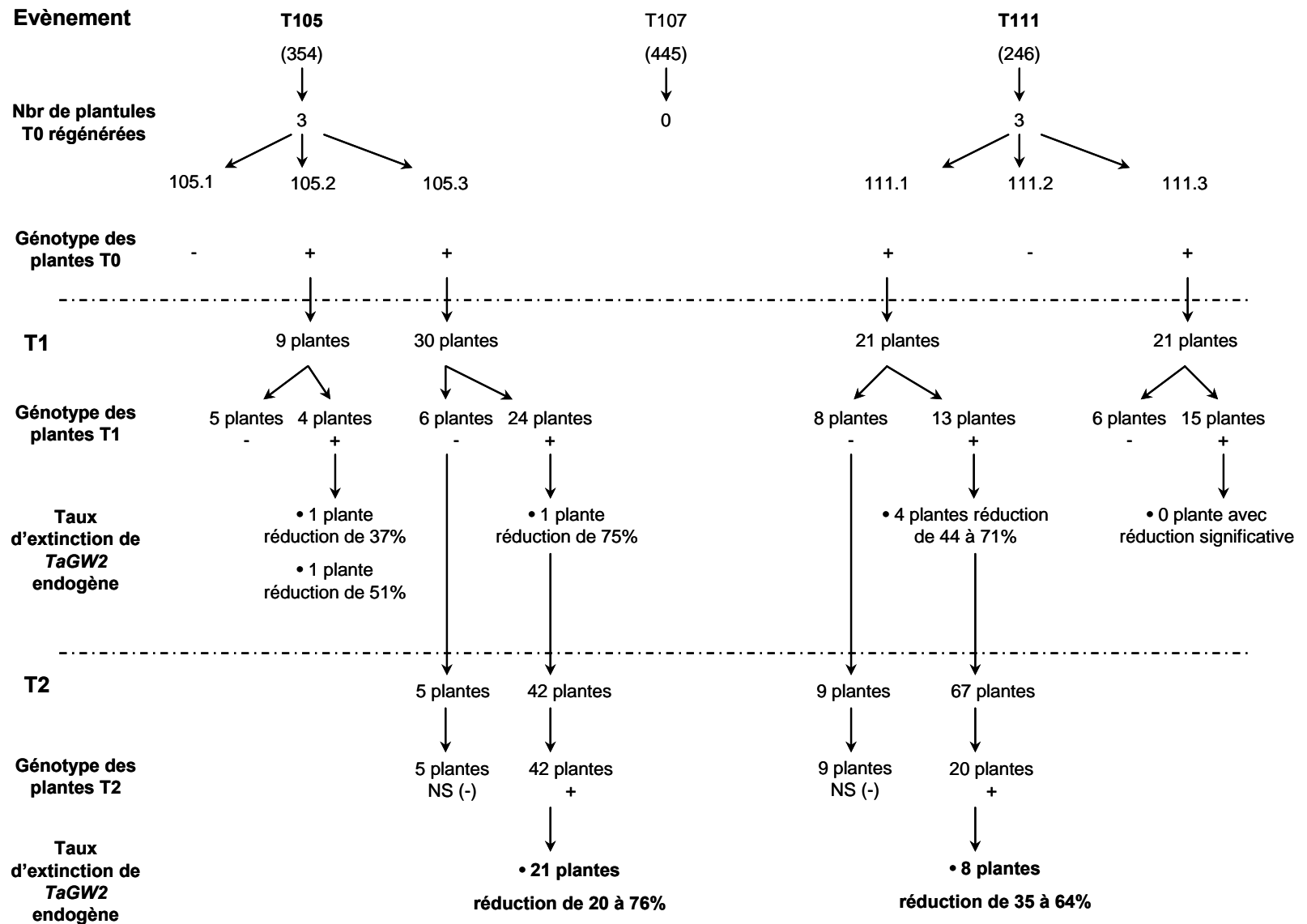


Figure 79 : Schéma du protocole expérimental de génération des lignées transgéniques RNAi T105 et T111. Le nombre d'embryons immatures bombardés au cours de chaque évènement de transformation est indiqué entre parenthèses. NS, nul-ségrégants

de cette séquence génomique prédit que cette séquence code pour une protéine putative d'un rétrotransposon de type Ty3-gypsy.

Dans un second temps, une analyse blastn dans le set de séquences génomiques « 454 » non-assemblées et dans le set de séquences génomiques assemblées, de la base de données CerealdB (www.cerealsdb.uk.net), a permis l'identification respective de 69 et 13 séquences présentant une identité de séquence continue de 20 nt avec la séquence partielle de *TaGW2-A*. Nous avons alors procédé à une analyse blastx dans NCBI pour chacune de ces séquences. Parmi les 13 séquences génomiques assemblées, 11 séquences sont prédites pour coder une E3 RING putative et présentent une homologie de séquence en acides aminés avec la protéine Yrg1 chez l'orge. Par conséquent, ces 11 séquences peuvent correspondre à *TaGW2*. Les 2 autres séquences restantes ne présentent des similarités avec aucune autre séquence connue.

Collectivement, ces analyses dans les bases de données de l'URGI et CerealdB indiquent que la cible de la construction RNAi *TaGW2* est probablement limitée aux copies homéologues de *TaGW2*. Cependant, nous ne pouvons pas totalement exclure l'extinction par cette cassette RNAi d'autres gènes que les copies *TaGW2*, du fait du séquençage incomplet à ce jour du génome du blé tendre.

1.2. Caractérisation moléculaire des lignées transgéniques RNAi *TaGW2*.

Trois séries (T105, T107 et T111) de 354, 445 et 246 embryons immatures de *T.aestivum*, variété Récital, ont été transformées par biolistique, par co-bombardement de la cassette RNAi *TaGW2* et d'une cassette de sélection comportant le gène de la phospho-mannose isomérase (PMI). Sur les 1045 embryons bombardés, 6 plantes ont pu être finalement régénérées après callogénèse puis organogénèse : 3 plantes issues de l'évènement de transformation T105 et 3 plantes issues de l'évènement T111 (Figure 79). Aucune plante n'a pu être régénérée à l'issue de l'évènement T107. Les plantules régénérées possèdent le gène PMI puisqu'elles ont été obtenues sur des milieux sélectifs pour la PMI. Cependant, le gène de sélection PMI n'appartenant pas physiquement à la cassette RNAi *TaGW2*, les plantules ont très bien pu intégrer la cassette de sélection sans intégrer la cassette RNAi *TaGW2*. Les plantules ont donc été génotypées pour la présence de la cassette RNAi *TaGW2*, par PCR sur ADN génomique, avec le couple d'amorces PDK-intron. Sur les 6 plantules, une plantule T105 (105.1) et une plantule T111 (111.2) ne présentaient pas de cassette RNAi et ont été

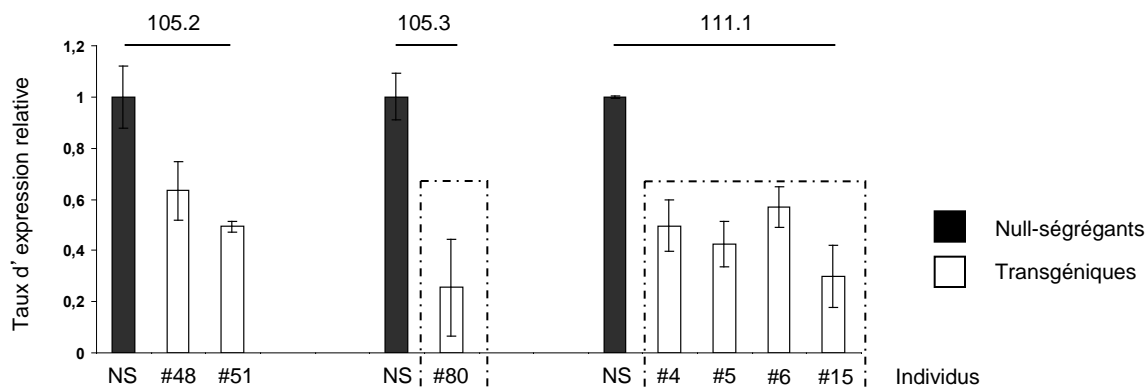


Figure 80 : Expression endogène relative de *TaGW2*, toutes copies confondues, dans des plantes T1 issues des plantes-mère 105.2, 105.3 et 111.1, par rapport à des plantes contrôle nulle-ségrégantes (NS). NS issues des plantes-mère 105.2, 105.3 et 111.1 : N=11 ; 3 ; et 3, respectivement. Les données ont été générées par PCR quantitative et sont des moyennes de triplicats de deux réplicats techniques \pm SD. *Ta54227* a été utilisé comme gène de référence. Les plantes auto-pollinisées pour générer la F2 des événements T105 et T111 sont encadrées en pointillés.

Tableau 12 : Expression endogène relative de *TaGW2*, toutes copies confondues, dans des plantes T2 transgéniques et nulle-ségrégantes des événements T105 (A) et T111 (B). Les données ont été générées par PCR quantitative et sont des moyennes de triplicats de deux réplicats techniques \pm SD. *Ta54227* a été utilisé comme gène de référence.

A T2 105		B T2 111	
Individu	Expression résiduelle relative de <i>TaGW2</i>	Individu	Expression résiduelle relative de <i>TaGW2</i>
Null-ségrégants (N=5)	1 \pm 0,09	Null-ségrégants (N=9)	1 \pm 0,11
105#1	0,76 \pm 0,02	111#1	0,45 \pm 0,15
105#2	0,25 \pm 0,06	111#2	0,66 \pm 0,14
105#3	0,67 \pm 0,02	111#3	0,60 \pm 0,02
105#4	0,67 \pm 0,03	111#4	0,63 \pm 0,09
105#5	0,59 \pm 0,16	111#5	0,62 \pm 0,18
105#6	0,64 \pm 0,19	111#6	0,56 \pm 0,01
105#7	0,73 \pm 0,11	111#7	0,75 \pm 0,03
105#8	0,67 \pm 0,00	111#8	0,36 \pm 0,06
105#9	0,49 \pm 0,25		
105#10	0,70 \pm 0,12		
105#11	0,69 \pm 0,06		
105#12	0,71 \pm 0,07		
105#13	0,69 \pm 0,12		
105#14	0,77 \pm 0,04		
105#15	0,74 \pm 0,00		
105#16	0,61 \pm 0,10		
105#17	0,64 \pm 0,14		
105#18	0,44 \pm 0,16		
105#19	0,80 \pm 0,06		
105#20	0,51 \pm 0,02		
105#21	0,70 \pm 0,10		

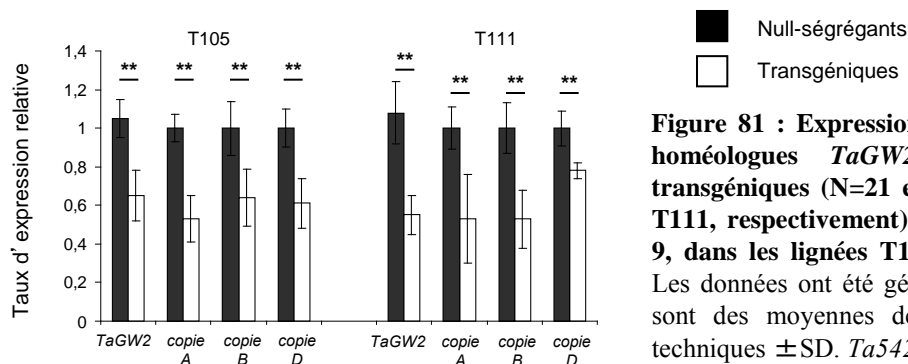


Figure 81 : Expression endogène relative des trois homologues *TaGW2* dans des plantes T2 transgéniques (N=21 et 8, dans les lignées T105 et T111, respectivement) et nulle-ségrégantes (N=5 et 9, dans les lignées T105 et T111, respectivement). Les données ont été générées par PCR quantitative et sont des moyennes de triplicats de deux réplicats techniques \pm SD. *Ta54227* a été utilisé comme gène de référence. **, $p < 0,01$ (test de Scheffé).

éliminées. Les 4 plantules T0 restantes (T105.2, T105.3, T111.1 et T111.3) ont été auto-pollinisées par apposition de sac d'auto-fécondation sur les épis.

Les plantes T1 ont été génotypées pour la présence de la cassette RNAi par PCR sur ADN génomique et caractérisées par PCR quantitative afin d'estimer le taux d'extinction du transcrit endogène *TaGW2*, toutes copies confondues, par rapport au taux d'expression endogène de *TaGW2* dans des plantes 'contrôle' nulle-ségrégantes. Sur les 9 plantes T1 issues de la plante-mère T105.2, 4 plantes étaient positives, c'est-à-dire possédaient la cassette RNAi *TaGW2*, dont 2 plantes présentaient des taux de réduction de l'expression endogène de *TaGW2* de 37% et 51% (Figure 80). Sur les 24 plantes T1 issues de T105.3, 24 plantes étaient positives dont 1 seule présentait une réduction importante de l'expression de *TaGW2* (75% de réduction ; Figure 80). Cette plantule T1 a été auto-pollinisée afin de générer la génération F2 de l'évènement T105. Sur les 21 plantes T1 issues de T111.1, 13 plantes étaient positives dont 4 présentaient un taux de réduction de l'expression endogène de *TaGW2* allant de 44 à 71% (Figure 80). Sur les 21 plantes T1 issues de T111.3, 15 plantes étaient positives mais aucune ne présentait un taux significatif de réduction de l'expression endogène de *TaGW2*. Seules les plantes présentant les taux d'extinction les plus importants ont été auto-pollinisées afin de générer les générations F2 des évènements T105 et T111.

Les plantes T2 ont été génotypées pour la présence de la cassette RNAi et caractérisées par PCR quantitative. Sur les 42 plantes T2 de l'évènement T105, toutes les plantes étaient positives dont 21 présentaient un taux d'extinction de l'expression endogène de *TaGW2* allant de 20 à 76%, par rapport aux plantes 'contrôle' (Tableau 12, A). Sur les 67 plantes T2 de l'évènement T111, 20 plantes étaient positives dont 8 plantes présentaient un taux d'extinction de *TaGW2* allant de 35 à 64%, par rapport aux plantes 'contrôle' (Tableau 12, B).

Le taux d'extinction de chacune des 3 copies homéologues de *TaGW2* a été estimée par PCR quantitative, avec des couples d'amorces copie-spécifique, dans les 21 et 8 plantes T2 des évènements T105 et T111, respectivement. Dans les 2 lignées transgéniques, nous avons observé une réduction significative de l'expression de chacune des copies *TaGW2-A*, *TaGW2-B* et *TaGW2-D* (Figure 81). L'expression de *TaGW2-A* et *-B* a été réduite de près de 45% dans les 2 lignées, T105 et T111 ; tandis que l'expression de *TaGW2-D* a été réduite de 39% et 22% dans la lignée T105 et T111, respectivement.

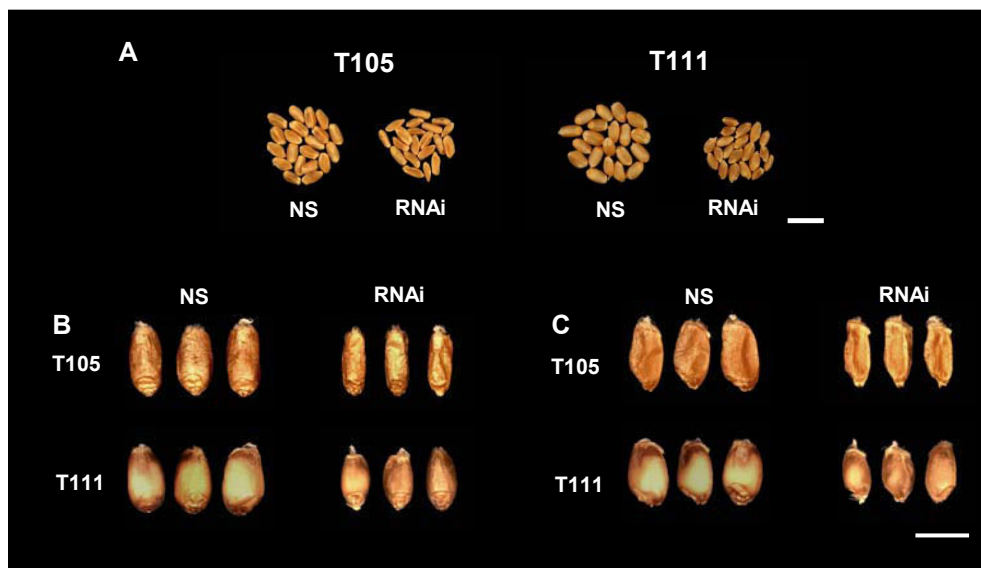


Figure 82 : Phénotypes des grains des plantes T2 des lignées transgéniques RNAi T105 et T111, comparés à des grains de plantes contrôle T2 null-ségrégantes (NS). (A) Vingt grains issus d'un même épi de plantes transgéniques et nulle-ségrégantes des 2 événements de transgénèse. Barre, 1 cm. (B-C) Vues dorsale (B) et latérale (C) de grains basaux d'épillets centraux d'épi du maître-brin. Barre, 5 mm.

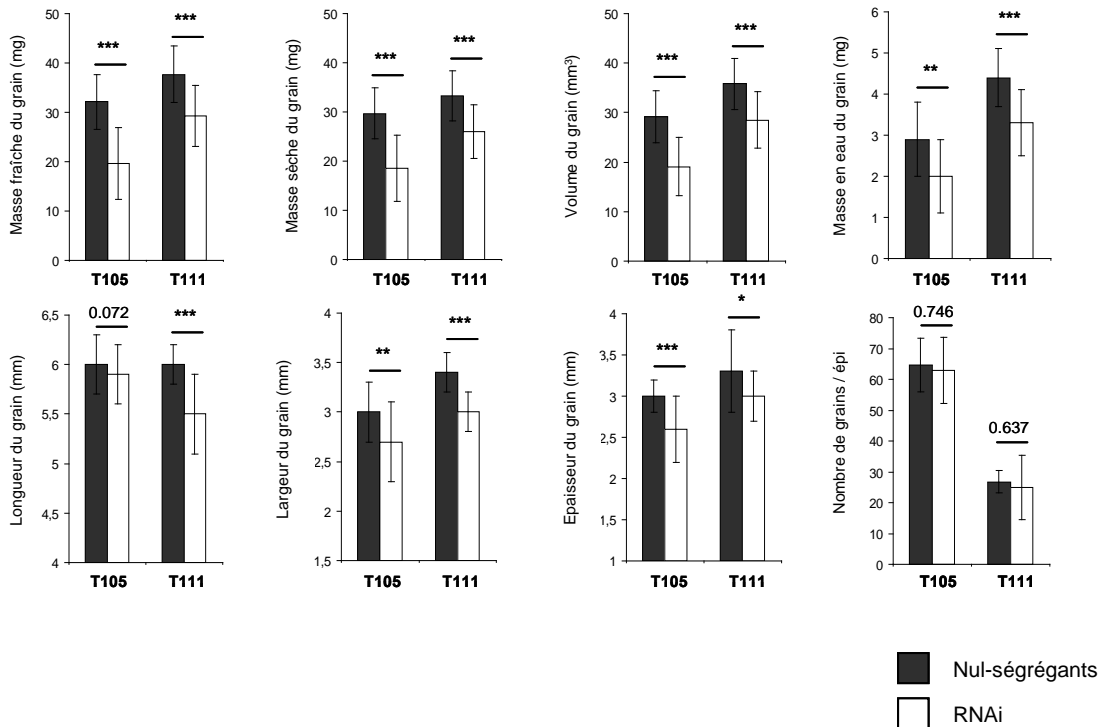


Figure 83 : Caractérisation phénotypique des plantes T2 des lignées RNAi T105 et T111. Le nombre de plantes T2 transgéniques et nulle-ségrégantes est n=21 et n=5, respectivement, pour l'évènement T105, et n=8 et n=9, respectivement, pour l'évènement T111. Les valeurs sont des moyennes \pm SD. Tous les paramètres, excepté le nombre de cellules de l'albumen, ont été mesurés sur des grains récoltés à maturité (1000° C Jours). Le nombre de cellules de l'albumen a été compté sur des grains récoltés à 260° C Jours. La date de floraison est exprimé en date julienne de l'année. Le nombre de jours entre le semis et la floraison étaient de 144 et 126 jours pour T105 et T111, respectivement. *, p<0,05 ; **, p<0,01 ; ***, p<0,001.

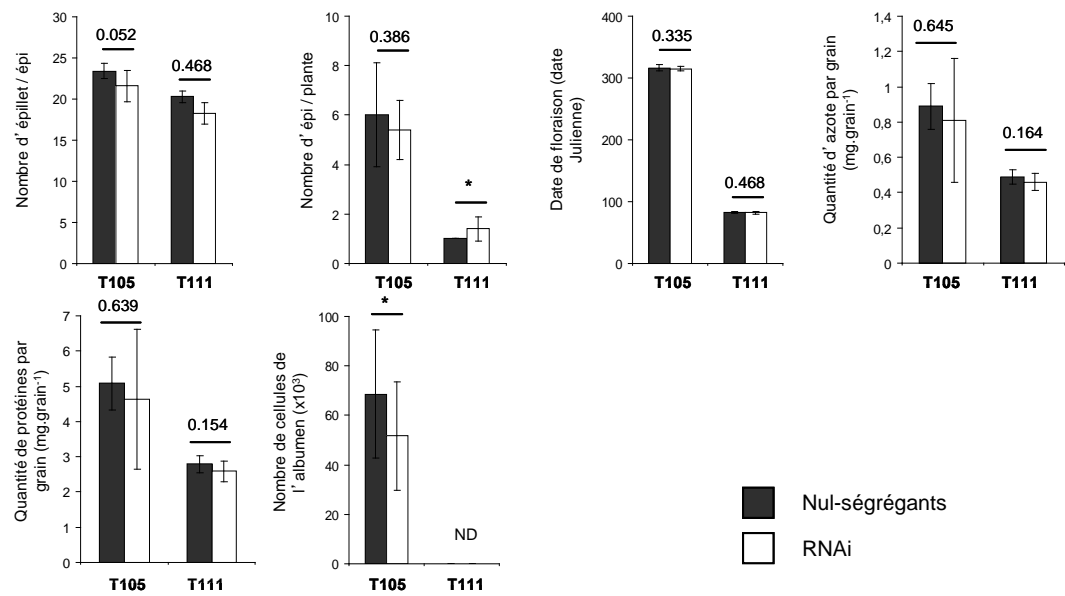
2. Phénotypage des lignées RNAi TaGW2

Afin de déterminer si *TaGW2* impacte le rendement en grain chez le blé tendre tout comme son homologue *GW2* chez le riz, les plantes T2 ont été phénotypées à maturité physiologique pour des composantes du rendement et notamment des caractères de la taille du grain. Les caractères suivants ont été mesurés sur l'ensemble des plantes transgéniques présentant une extinction de *TaGW2* et les plantes nulle-ségrégantes : le nombre d'épis par plante ; le nombre d'épillets par épi ; le nombre de grains par épi ; la date de floraison ; les masses fraîches et sèches du grain ; le volume, la largeur, l'épaisseur et la longueur du grain ; la masse en eau du grain à maturité ; ainsi que les quantités d'azote et de protéines par grain. Comme il est établi que la position de l'épillet le long de l'épi impacte le poids final et le contenu en azote des grains inclus dans cet épillet (Calderini et Ortiz-Monasterio, 2003 ; Andersson *et al*, 2004), ainsi, afin de pouvoir comparer le plus justement possible les valeurs de chacun des caractères entre les lignées transgéniques et contrôle, toutes les mesures ont été effectuées sur les deux grains basaux des deux épillets centraux de l'épi du maître-brin de chaque plante. Au sein de chaque lignée transgénique, une analyse de variance a été réalisée sur les valeurs obtenues à l'aide du logiciel SAS/STAT®, et les moyennes ont été comparées par un test de Student-Newman-Keuls.

Un effet significatif de la sous-expression des copies *TaGW2* sur la taille et le poids du grain a été observé dans les lignées transgéniques par rapport aux plantes contrôle (Figure 82). La masse fraîche du grain a été réduite de 38,6% et 22,2% dans les plantes transgéniques comparées aux plantes contrôle, dans les lignées T105 et T111, respectivement (Figure 83) ; la masse sèche du grain réduite de 37,3% et 21,9% ; et la masse en eau du grain réduite de 31% et 25%. Les plantes transgéniques présentent également une réduction de 24,6% et 20,4% du volume du grain ; de 10% et 11,8% de la largeur du grain ; et de 13,3% et 9,1% de l'épaisseur du grain, par rapport aux plantes contrôle, dans les lignées T105 et T111, respectivement. Nous n'avons observé aucune différence significative entre les plantes transgéniques et les plantes contrôle pour le nombre d'épillets par épi, le nombre de grains par épi, la date de floraison et les quantités d'azote et de protéines par grain, dans les 2 lignées T105 et T111. Des différences significatives dans la longueur du grain et le nombre d'épis par plante entre les plantes transgéniques et les plantes contrôle ont été observées uniquement dans la lignée T111.

La taille finale du grain de blé tendre est pratiquement établie à la fin de la phase de division du développement du grain, estimée vers 260°CJours dans la variété Récital ; et le nombre de cellules de l'albumen est l'un des facteurs déterminants chez le blé tendre. Afin de

Figure 83 (suite)



déterminer si *TaGW2* impacte la taille finale du grain de blé tendre, les grains basaux d'épillets centraux d'épis de talles secondaires ont été récoltés à 260°CJours, dans des plantes transgéniques et des plantes contrôle de la lignée T105 ; et le nombre de cellules de l'albumen a été compté. Une réduction significative de 25% du nombre de cellules de l'albumen a été observée dans les grains de plantes transgéniques comparés aux grains de plantes contrôle, de la lignée T105 (Figure 83). Malheureusement, ce prélèvement de grains à 260°CJours n'a pu être effectué sur des plantes de la lignée T111 du fait du très faible nombre d'épis produits par les plantes transgéniques comme les plantes contrôles dans cette lignée. Le nombre d'épis est le reflet de la capacité de tallage d'une plante. Il a été montré chez le blé que la capacité de tallage pouvait répondre à des différences dans la longueur du jour ; des blés ayant poussés en jours courts produisant plus de talles. Les plantes T2 de chaque lignée transgénique ont été conduites en serre à des périodes différentes de l'année : les plantes T2 de l'évènement T105 ont poussé de juin 2010 à novembre 2011 soit en jours décroissants, et les plantes T2 de l'évènement T111 de novembre 2010 à mai 2011, soit en jours croissants. Néanmoins, la variété Récital est connue pour être une variété insensible à la photopériode et une différence dans la longueur du jour ne semble pas être une explication plausible à la différence du nombre d'épis produits entre l'ensemble des plantes T105 et l'ensemble des plantes T111. De plus, un complément d'éclairage a été apporté tout au long de la croissance des plantes T2 en serre « transgénique », afin de compenser ces différences dans la photopériode. La capacité de tallage d'une plante est également affectée par d'autres paramètres environnementaux tels que le stress hydrique, la densité de semis, la disponibilité en nutriments et la remobilisation des assimilats. Les plantes des 2 évènements de transgénèse ont été traitées de manière identique quant au terreau utilisé, les conditions de température dans la serre transgénique et l'arrosage. Cependant, du fait de limitations de place en serre transgénique, les plantes T2 de l'évènement T111 ont été repiquées en pots de 1L contre des pots de 3L pour les plantes T2 de l'évènement T105. Cette différence dans le volume des pots utilisés pourrait mimer une densité de semis différente, reportée comme influençant fortement la capacité de tallage, et pourrait être à l'origine de cette différence du nombre d'épis entre les plantes T2 de la lignée T105 et de la lignée T111.

3. Production de lignées de surexpression *TaGW2-A*

Afin de surexprimer *TaGW2-A* dans le grain de blé uniquement, une cassette a été construite comportant l'ADNc complet de *TaGW2-A* en aval d'un promoteur fort graine-

spécifique. Quatre séries (T117, T118, T119 et T120) de 326, 308, 302 et 291 embryons immatures de *T.aestivum*, variété Bobwhite, ont été transformées par biolistique, par bombardement de la cassette de surexpression Bx7-TaGW2A-PMI. Sur les 1227 embryons bombardés, 22 plantes ont pu être finalement régénérées après callogénèse puis organogénèse : 3 plantes issues de l'évènement de transformation T117, 1 plante de l'évènement T118, 1 plante de l'évènement T119 et 17 plantes issues de l'évènement T120. Dans l'attente d'un génotypage de ces plantes par Southern-blot afin de déterminer le nombre de cassettes insérées, les plantes ont été auto-pollinisées par apposition de sachets d'auto-fécondation pour la production des grains T1.

Conclusions et Discussions du Chapitre II

En conclusion de ces travaux de transgénèse, nous avons montré que la sous-expression du gène *TaGW2* chez le blé entraîne des réductions significatives des masses finales fraîche et sèche du grain ; de la quantité d'eau du grain à maturité ; et du volume, de la largeur et de l'épaisseur du grain. Le gène *TaGW2* est donc un régulateur positif de la taille finale et du poids du grain, chez le blé tendre.

Aucune différence significative n'a été observée dans le nombre de grains par épi entre les plantes transgéniques et contrôle. Etant donné qu'une corrélation négative est souvent observée entre le nombre de grains par épi et la taille des grains (Fisher *et al*, 1977), la réduction significative du poids et des dimensions des grains dans les plantes transgéniques ne peut par conséquent pas être attribuée à un nombre plus important de grains par épi dans ces plantes. De plus, aucunes différences significatives dans le nombre d'épillets par épi, le nombre d'épis par plante (dans la lignée T105) et la date d'anthèse n'ont été relevées entre les plantes transgéniques et contrôles. Ceci suggère que l'expression du transgène n'a aucun effet sur le développement végétatif des plantes transgéniques et que par conséquent un effet intrinsèque direct du transgène sur le développement du grain est plausible.

Bien que des résultats similaires aient été obtenus dans les deux lignées transgéniques T105 et T111, de manière intéressante, l'effet de la sous-expression de *TaGW2* sur les composantes du rendement en grain était toujours moins important dans la lignée T111 que dans la lignée T105. Dans les deux lignées transgéniques, nous avons observé une réduction des copies A et B de *TaGW2* à des taux similaires (Figure 81). Cependant, l'accumulation des transcrits *TaGW2-D* a été réduite en moyenne deux fois moins dans les plantes transgéniques de la lignée T111 que dans les plantes transgéniques de la lignée T105. Cette différence dans l'extinction de la copie *TaGW2-D* entre les deux lignées transgéniques pourrait par conséquent expliquer la différence d'amplitude observée de l'effet du transgène entre les lignées T105 et T111. Ce résultat suggère que les différentes copies homéologues contribuent à ce caractère quantitatif, et que la modulation du taux d'expression d'une seule des copies homéologues de *TaGW2* peut quantitativement impacter des composantes du rendement.

Quelques rappels sur la mise en place du potentiel de poids du grain chez les céréales

La détermination du potentiel de taille finale et de poids du grain constitue un enjeu économique important et fait l'objet de nombreuses études chez le blé et d'autres céréales depuis une cinquantaine d'années. Comme pour tout organe en développement, la mise en

place du potentiel du grain est complexe et semble reposer sur de nombreux processus fortement inter-dépendants. Nous avons essayé ici d'en dresser les grandes lignes et tenté d'intégrer *TaGW2* au sein de tous ces processus interconnectés.

Le potentiel de poids du grain est largement déterminé en premier lieu par le volume du grain puis sa capacité de remplissage en réserves amylacée et protéique. Contrairement à la grande majorité des graines, dans lesquelles la taille de la graine est largement déterminée par la taille de l'embryon (le tissu nourricier étant digéré au cours de la croissance de l'embryon), la taille du caryopse des céréales est conditionnée par le volume du tissu nourricier du grain, l'albumen, et non par la taille de l'embryon. Le volume du grain résulte donc du nombre et de la taille des cellules de l'albumen et est déterminé par les trois dimensions longueur, largeur et épaisseur. En étudiant les composantes génétiques déterminant la taille et la forme du grain de blé, Gegas et ses collaborateurs (2010) ont montré que les dimensions largeur et longueur étaient fortement découplées dans le programme de développement du grain ; la longueur étant pratiquement uniquement une composante génétique de la forme du grain, et la largeur une composante de la taille et du poids du grain. En effet, les auteurs ont montré dans 6 populations et 9 environnements qu'il y a entre autre une forte corrélation positive entre poids de 1000-grains, surface de la section transversale médiane du grain et largeur ; mais une faible voire pas de corrélation entre poids de 1000-grains et longueur. Cependant, de manière contrastée, Lizana et ses collaborateurs (2010) ont montré dans leur étude que la longueur du grain était la dimension présentant la meilleure corrélation avec le poids final du grain et le volume du grain. La croissance volumique du grain de blé a lieu majoritairement au cours de la phase de division. Chez le blé, la longueur est fixée précocement, à la fin de la phase de division, et 90% de la longueur totale du grain est déjà atteinte à 140°CJours (Rogers et Quatrano, 1983). Bien qu'elles soient majoritairement établies à la phase de division, les dimensions largeur et épaisseur du grain n'atteignent leurs dimensions maximales qu'aux alentours de 650°CJours, dans la variété Récital, et sont donc souvent considérées également comme des marqueurs du remplissage du grain de blé (C. Girousse, communication personnelle).

La division cellulaire dans l'albumen s'arrête vers 240-260°CJours. A l'heure actuelle, deux thèses s'affrontent quant au tissu déterminant le potentiel de la taille de l'albumen chez le blé. Jusqu'à présent, il était communément admis que la taille de l'albumen était conditionnée par le tissu lui-même, par un contrôle de la division cellulaire intrinsèque aux cellules de l'albumen. Des études récentes suggèrent que le contrôle du nombre de cellules de

l'albumen se ferait plutôt de manière indirecte : dans leur modèle, ce sont les téguments de la graine qui délimiteraient le volume potentiel de l'albumen (Calderini *et al*, 1999a, b ; Calderini et Reynolds, 2000 ; Garcia *et al*, 2005 ; Ugarte *et al*, 2007). Ainsi, la division cellulaire des cellules de l'albumen se ferait jusqu'à occuper entièrement le volume précédemment établi par les téguments. Cette hypothèse est notamment plausible du fait qu'un modèle similaire de croissance expansive conditionnée par la croissance des tissus extérieurs a été reporté pour d'autres organes végétatifs (Kutschera et Niklas, 2007). Pendant cette période de division cellulaire, la croissance volumétrique du grain est largement déterminée par l'accumulation rapide d'eau dans le grain (Schnyder et Baum, 1992). L'arrêt de la division cellulaire coïncide notamment avec l'arrêt d'accumulation d'eau dans le grain (Nadaud *et al*, 2012). De plus, le contenu en eau jouerait un rôle dans l'élongation cellulaire puisque la taille finale des cellules de l'albumen serait en partie influencée par l'accumulation en eau (Egli, 1998).

La masse sèche finale du grain, qui résulte de l'accumulation de la matière sèche (réserves amylacées et protéiques du grain), est largement déterminée par le nombre de cellules de l'albumen et leur taille, au moment de la phase de division du grain (Hoshikawa, 1962 ; Brocklehurst, 1977 ; Vilhar *et al*, 2002). Au cours de la phase de remplissage du grain, la masse sèche du grain augmente rapidement avec les protéines s'accumulant majoritairement de 14 à 21 JAA (soit environ 280°CJours à 420°CJours) et l'amidon majoritairement de 14 à 35 JAA (soit environ 280°CJours à 700°CJours). Dans leur étude, Schnyder et Baum (1992) montrent que la masse sèche du grain à maturité est très faiblement corrélée à la durée du remplissage, mais fortement corrélée à la vitesse de remplissage en matière sèche. Les auteurs indiquent également que la masse sèche finale du grain est atteinte lorsque le contenu relatif en matière sèche est de 0,55g/g de masse fraîche. Le statut hydrique du grain et la dynamique d'accumulation de matière sèche sont étroitement liés (Schnyder et Baum, 1992 ; Calderini et Reynolds, 2000 ; Lizana *et al*, 2010). Il existe une relation positive et linéaire entre le contenu maximal en eau du grain et la masse sèche du grain à maturité (Schnyder et Baum, 1992). Le contenu maximal en eau du grain est atteint précocement au cours du développement du grain, lorsque un tiers de la masse sèche maximale est atteint (Schnyder et Baum, 1992). De manière intéressante, la masse sèche maximale du grain de blé peut être prédite de manière relativement fiable dès que les grains ont pratiquement atteint leur contenu maximal en eau. De manière contrastée, une prédiction à des stades antérieurs est

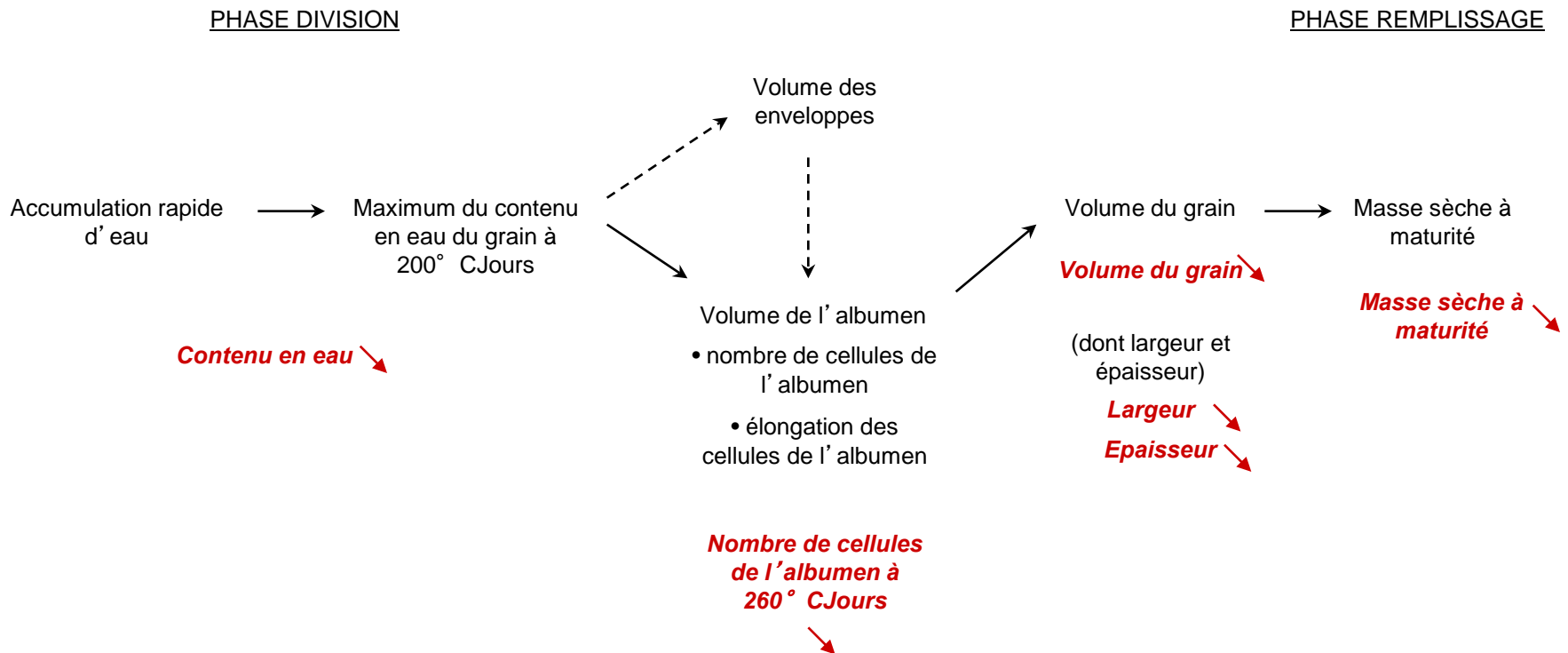


Figure 84 : Correspondance entre les phénotypes des grains des plantes transgéniques RNAi *TaGW2* et les processus cellulaires déterminant le potentiel de poids du grain lors de la phase de division. Les phénotypes observés dans les plantes transgéniques sont indiqués en rouge.

beaucoup moins fiable, indiquant que la capacité puits des grains ne peut pas être établie quand le contenu en eau du grain est en train d'augmenter rapidement. Schnyder et Baum (1992) suggèrent que la restriction volumique, préalablement établie par la dynamique hydrique du grain pendant la phase de division, est probablement impliquée dans la détermination de la capacité puits du grain. Les auteurs ont montré que le contenu en eau du grain augmentait rapidement lorsque la quantité relative de matière sèche était de 0,2 à 0,3g/g de masse fraîche. Le contenu en eau se stabilise alors (plateau hydrique) pour des valeurs de quantités relatives de matière sèche de 0,33 à 0,55g/g de masse fraîche du grain. La fin d'accumulation d'eau coïncide notamment avec la fin d'accumulation de la matière sèche. Schnyder et Baum (1992) mentionnent que l'arrêt d'accumulation de la matière sèche correspondrait à un arrêt de la synthèse d'amidon, la quantité d'amidon ayant atteint une densité critique. Plusieurs études ont proposé une même explication à l'arrêt de ce processus de remplissage (Zee et O'Brian, 1970 ; Zee, 1975 ; Sofield *et al*, 1977b ; Caley *et al*, 1990). Le dépôt de composés lipidiques dans la « couche pigmentée » (leur terme, soit par déduction la couche à aleurone) constituerait une barrière à l'entrée d'eau dans le grain, créant alors une diminution du potentiel hydrique dans l'albumen. L'enzyme ADPglucose pyrophosphorylase, et probablement d'autres enzymes impliquées dans le métabolisme carboné dans l'albumen, sensible au déficit hydrique, voit alors son activité chuter brutalement.

TaGW2 dans le développement du grain de blé tendre

TaGW2 au cours de la phase de division du grain de blé tendre

L'eau s'accumule rapidement dans le grain au cours de la phase de division, atteignant un contenu maximal vers 200°Cjours qui conditionne le volume du grain (Figure 84), par le biais du volume des enveloppes du grain ou directement le volume de l'albumen lui-même, lui-même principalement déterminé par les dimensions « largeur » et « épaisseur », et par conséquent influence directement l'accumulation en matière sèche et donc la masse sèche du grain. De manière intéressante, dans les plantes transgéniques éteintes pour *TaGW2*, nous avons observé des réductions de la quantité totale d'eau dans le grain à maturité, du volume du grain dont des réductions de la largeur et de l'épaisseur, et de la masse sèche du grain (Figure 84). De plus, nous avons observé une réduction du nombre de cellules de l'albumen. Les profils d'accumulation des transcrits homéologues montrent une expression de *TaGW2-A* et *-D* au cours de la phase de division puis une chute de leur expression à la fin de cette phase (240°CJours), et sont cohérents avec une action de *TaGW2* au cours de la phase de division.

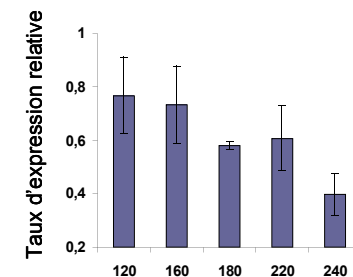
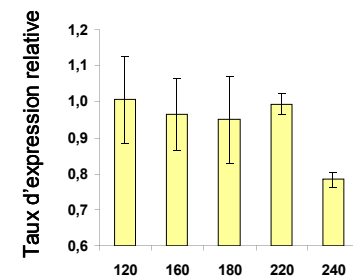
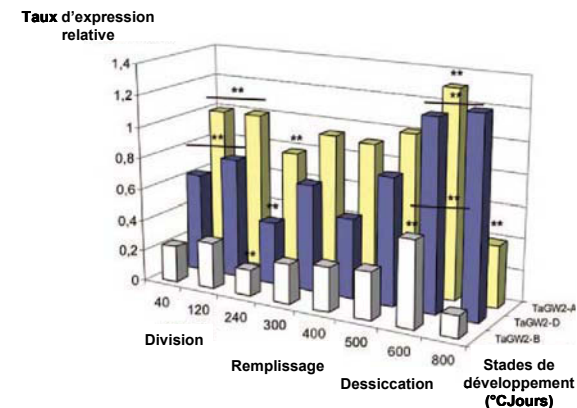
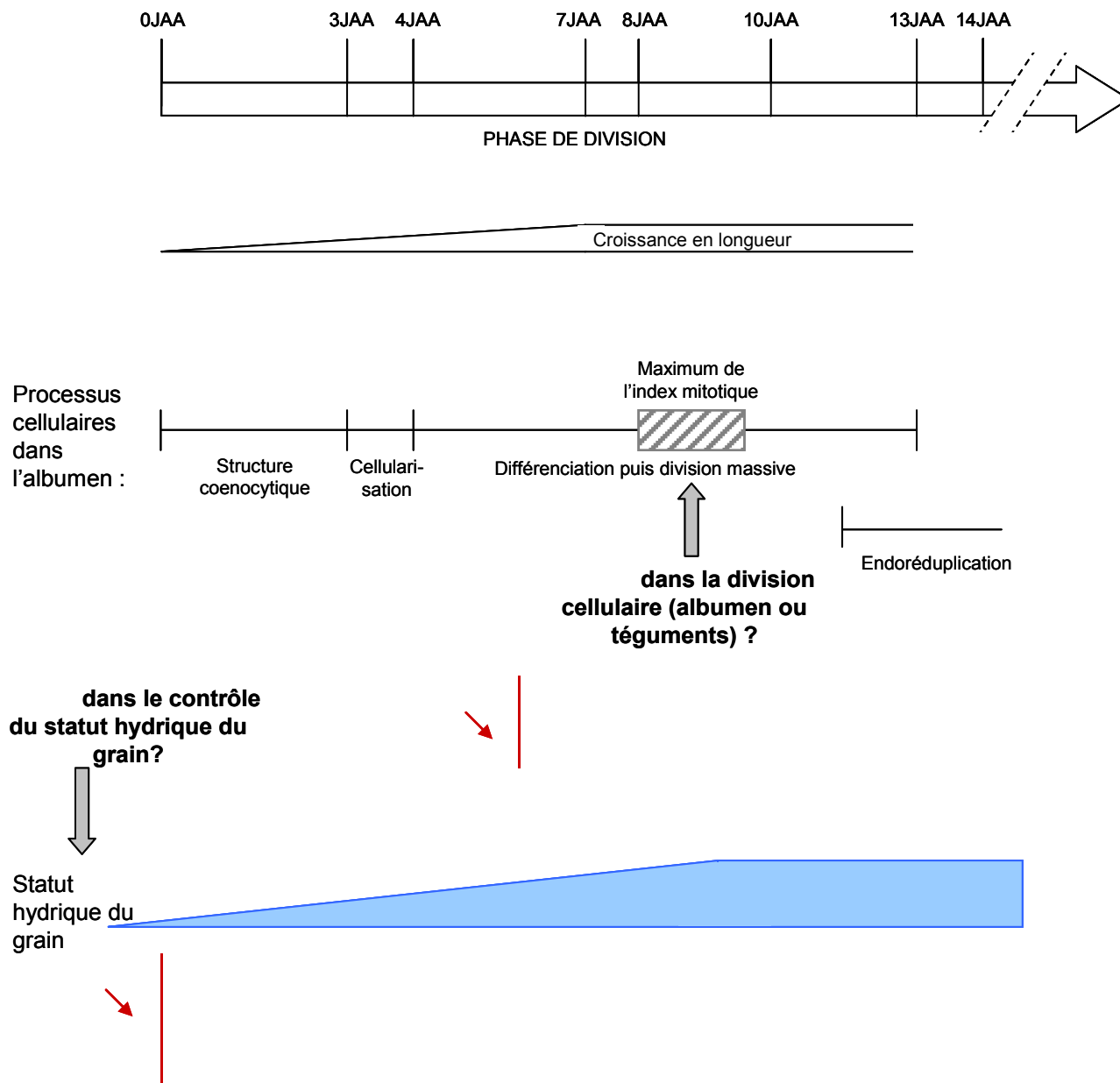


Figure 85 : Relation entre le profil d'accumulation en transcrit *TaGW2*, les phénotypes des grains transgéniques RNAi *TaGW2* et les processus cellulaires de détermination du potentiel de taille et de poids du grain

Plusieurs hypothèses peuvent être formulées quant au rôle de *TaGW2* au cours cette phase de développement (Figure 85). La première hypothèse est que *TaGW2* contrôle la dynamique hydrique dans le grain. En effet, tous les paramètres influencés par ce statut hydrique et mis en place au cours de la phase de division sont impactés par l’extinction de *TaGW2*. La deuxième hypothèse serait que *TaGW2* contrôle directement le processus de division cellulaire. Considérant les deux thèses actuelles sur la détermination du volume de l’albumen, *TaGW2* pourrait soit agir intrinsèquement dans l’albumen ou contrôler la division des cellules dans les téguments du grain. De manière intéressante, le profil d’accumulation des transcrits *TaGW2-A* et *-D* coïncide notamment avec le pic de l’index mitotique relevé dans l’albumen de céréales au cours de la phase de division (environ 160 à 200°CJours). En effet, le taux des transcrits *TaGW2-A* et *-D* à 120°CJours est maintenu jusqu’à 220°CJours puis chute à 240°CJours (Figure 85). Ceci pourrait constituer un argument en faveur de *TaGW2* comme régulateur de la division cellulaire dans l’albumen. Les expériences de localisation tissulaire de transcrit par hybridation *in situ*, actuellement mises en place au sein de l’équipe sur différents stades de développement de grain, pourraient permettre de savoir si *TaGW2* agit directement dans les cellules de l’albumen ou indirectement par une action dans les cellules des téguments. De plus, les cellules de transfert, différenciées dans l’albumen au moment de la phase de division, favorisent l’accumulation rapide de nutriments par l’albumen au cours de cette phase (Sabelli et Larkins, 2009) et il serait également intéressant de voir si le transcrit *TaGW2* s’accumule dans ces cellules au cours de la phase de division du grain.

Le phénomène d’endoréduplication constitue un des moyens d’augmentation du potentiel de la taille du caryopse (Sabelli, 2012). Il a été proposé notamment que l’endoréduplication permettrait d’augmenter la quantité de matrice afin de supporter des taux de transcription forts nécessaires à une expansion cellulaire rapide. Chez le maïs, la mise en place de cette phase a été décrite peu après 11 JAA (Sabelli et Larkins, 2009). Une détermination précise de la fenêtre temporelle de mise en place de ce processus cellulaire chez le blé tendre est nécessaire afin de déterminer si le profil d’accumulation des transcrits *TaGW2* peut coïncider avec ce processus cellulaire.

TaGW2 au cours de la phase de remplissage du grain de blé tendre

Au cours de la phase de remplissage, les réserves amylicée et protéique s’accumulent progressivement dans les cellules de l’albumen et la couche à aleurone. Dans notre étude, l’analyse transcriptomique des homéologues *TaGW2* au cours du développement du grain a révélé une expression progressive et maximale à 600°CJours des trois copies de *TaGW2*,

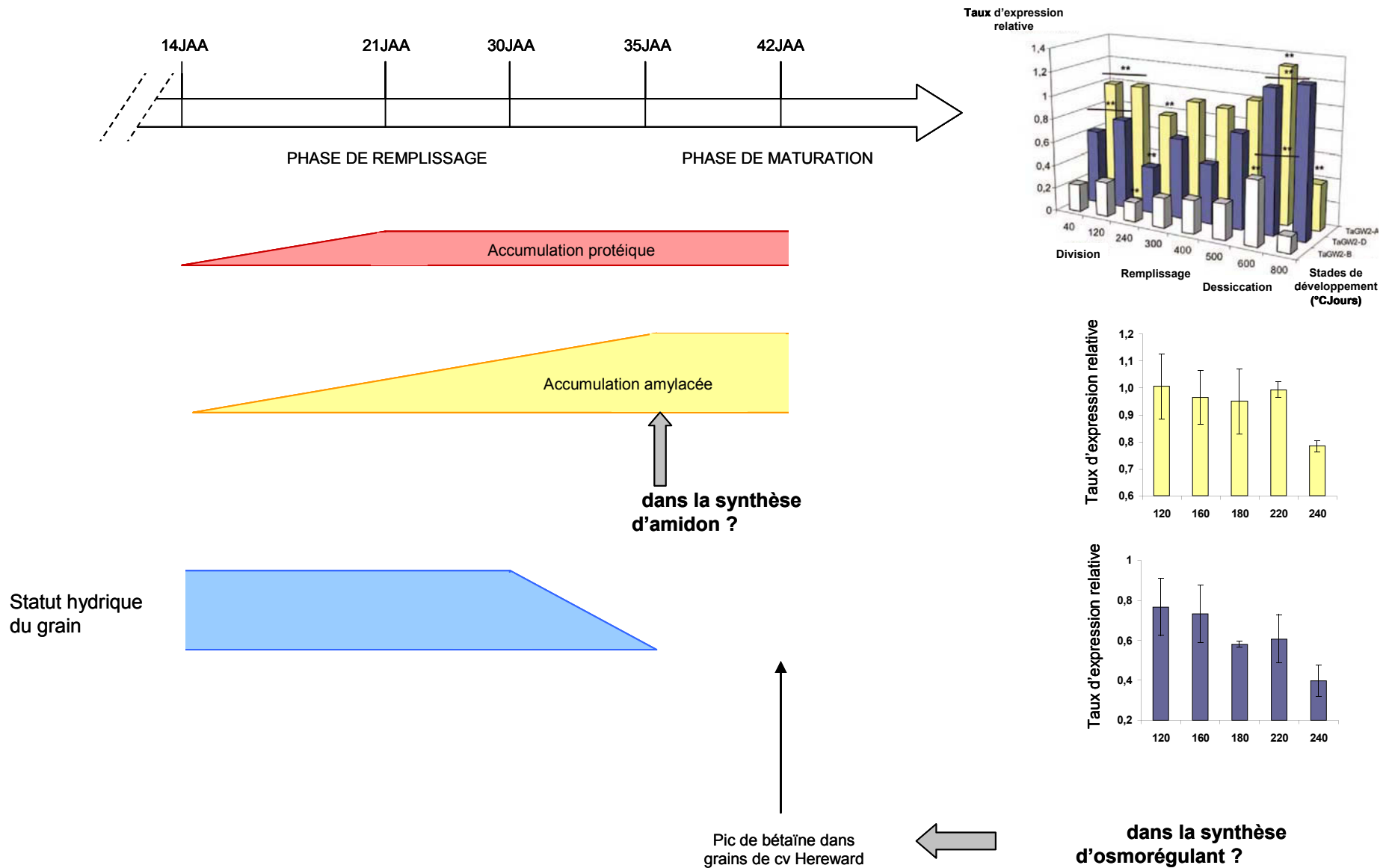


Figure 85 (Suite) : Les profils d'accumulation des transcrits *TaGW2-A*, *-B* et *-D* sont rappelés à droite de la figure. Les phénotypes observés dans les grains transgéniques sont indiqués en rouge

suggérant une action de *TaGW2* au milieu et la fin de la phase de remplissage. De plus, nous avons observé dans les plantes transgéniques éteintes pour *TaGW2* une réduction significative de la masse sèche du grain à maturité mais aucune différence dans la quantité d'azote et de protéines dans le grain par rapport aux plantes contrôles. Bien que le potentiel d'accumulation en matière sèche apparaisse être conditionné très tôt dans le développement par le statut hydrique, nous pouvons émettre l'hypothèse que *TaGW2* aurait également un effet sur l'accumulation de la matière sèche pendant la période de remplissage. De manière intéressante, le profil d'accumulation des transcrits *TaGW2* au cours de la phase de remplissage suit assez bien le profil d'accumulation en amidon. Nous proposons donc que *TaGW2* pourrait contrôler de manière positive la synthèse d'amidon. Contrairement aux résultats obtenus en génétique d'association pour *TaGW2-A*, *TaGW2* ne semblerait cependant pas influencer la synthèse en protéines de réserves au vu des résultats de transgénèse. Le profil d'accumulation des transcrits *TaGW2* au cours de la phase de remplissage semble cohérent avec cette observation puisqu'il suit mieux l'accumulation en réserve amylacée que protéique. Il est donc fort probable que les associations que nous avons détectées entre *TaGW2-A* et la quantité de protéines dans le grain résultent en fait d'un déséquilibre de liaison avec le polymorphisme causal situé dans un autre gène que *TaGW2-A*, comme nous l'avons présumé dans la discussion du chapitre I.

Au cours de la phase de remplissage, l'accumulation en eau se stabilise et le statut hydrique du grain commence à chuter brutalement vers environ 650°CJours. De manière intéressante, l'expression des copies *-A* et *-B* est corrélée au statut hydrique à la fin de la phase de remplissage puisqu'elle chute après 600°CJours, tandis qu'un taux élevé d'expression de *TaGW2-D* est maintenu. Cette observation peut renforcer l'idée que les copies A et B de *TaGW2* soient des régulateurs positifs du statut hydrique. *TaGW2-D* pourrait intervenir par exemple dans la synthèse d'osmorégulateurs pour la préparation du grain à la dormance. En effet, Shewry et ses collaborateurs (2012) ont détecté chez le blé le pic de synthèse de bétaine, un métabolite secondaire jouant le rôle d'osmorégulant chez d'autres espèces, vers 42JAA, soit 840°CJours.

L'ensemble de ces hypothèses sur la fonction de *TaGW2* a été bâti en s'aidant du profil d'accumulation des transcrits homéologues *TaGW2* et constitue une première approche. Des profils d'accumulation de transcrits n'étant pas systématiquement le reflet de

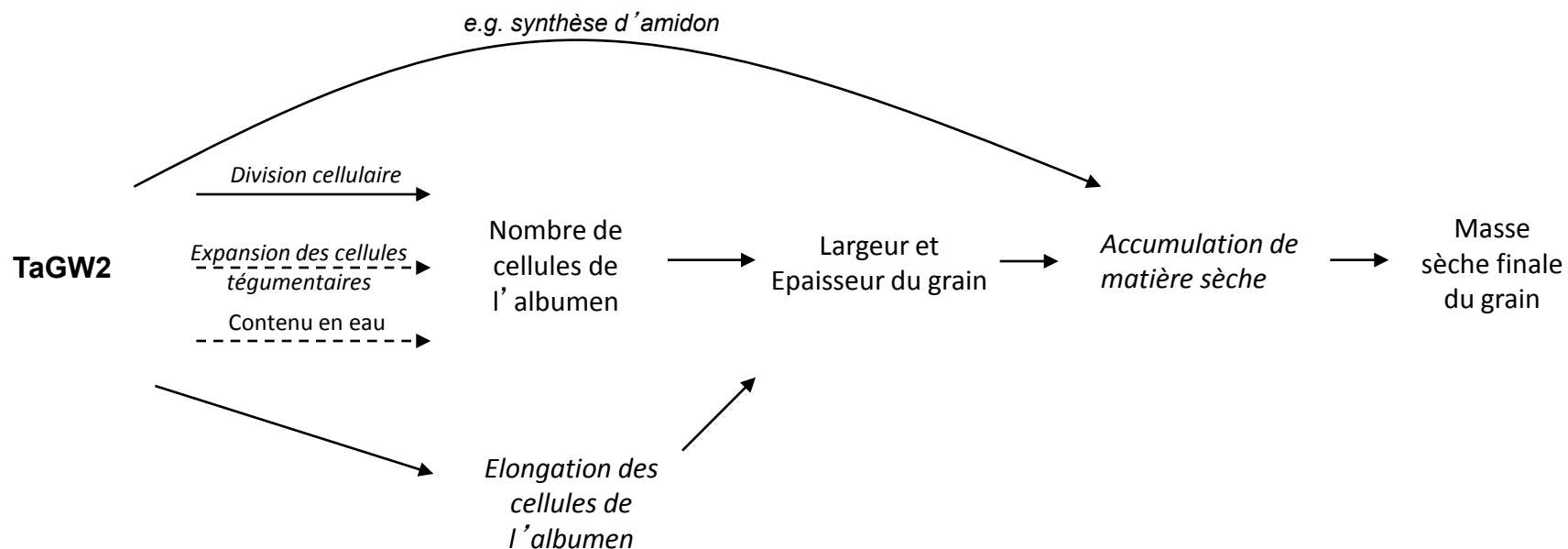


Figure 86 : Modèle proposé pour le rôle de *TaGW2* dans la régulation de la taille et du poids du grain. Les processus cellulaires sont indiqués en italiques. Les flèches en traits plein et en pointillés indiquent respectivement un effet direct et indirect de *TaGW2* sur le caractère ou processus.

l'accumulation de la molécule effective *in fine*, la protéine, ces hypothèses nécessitent bien entendu d'être confortées par l'analyse des profils d'accumulation protéique de *TaGW2* au cours du développement du grain.

Nous proposons le modèle suivant quant au rôle de *TaGW2* dans le développement du grain (Figure 86). Pour le moment, il est encore difficile de clairement identifier la part de chacun des homéologues *TaGW2* dans ces divers processus. Par conséquent, *TaGW2*, toutes copies confondues, sera considéré dans ce modèle. *TaGW2* régulerait positivement le nombre et la taille des cellules de l'albumen, influençant la largeur et l'épaisseur du grain, conduisant à une accumulation accrue de la matière sèche et par conséquent augmentant la masse sèche finale du grain. *TaGW2* pourrait contrôler le nombre de cellules de l'albumen directement en étant un régulateur positif de la division cellulaire dans l'albumen ou indirectement en établissant les restrictions volumétriques à la croissance de l'albumen en contrôlant positivement soit l'expansion des cellules des téguments ou l'accumulation en eau du grain au cours de la phase de division. Nous n'excluons pas que *TaGW2* puisse contrôler directement également l'accumulation de la matière sèche pendant la période de remplissage par une régulation positive de la synthèse de polymères tels que l'amidon par exemple.

Chapitre III :

Recherche de partenaires protéiques de TaGW2

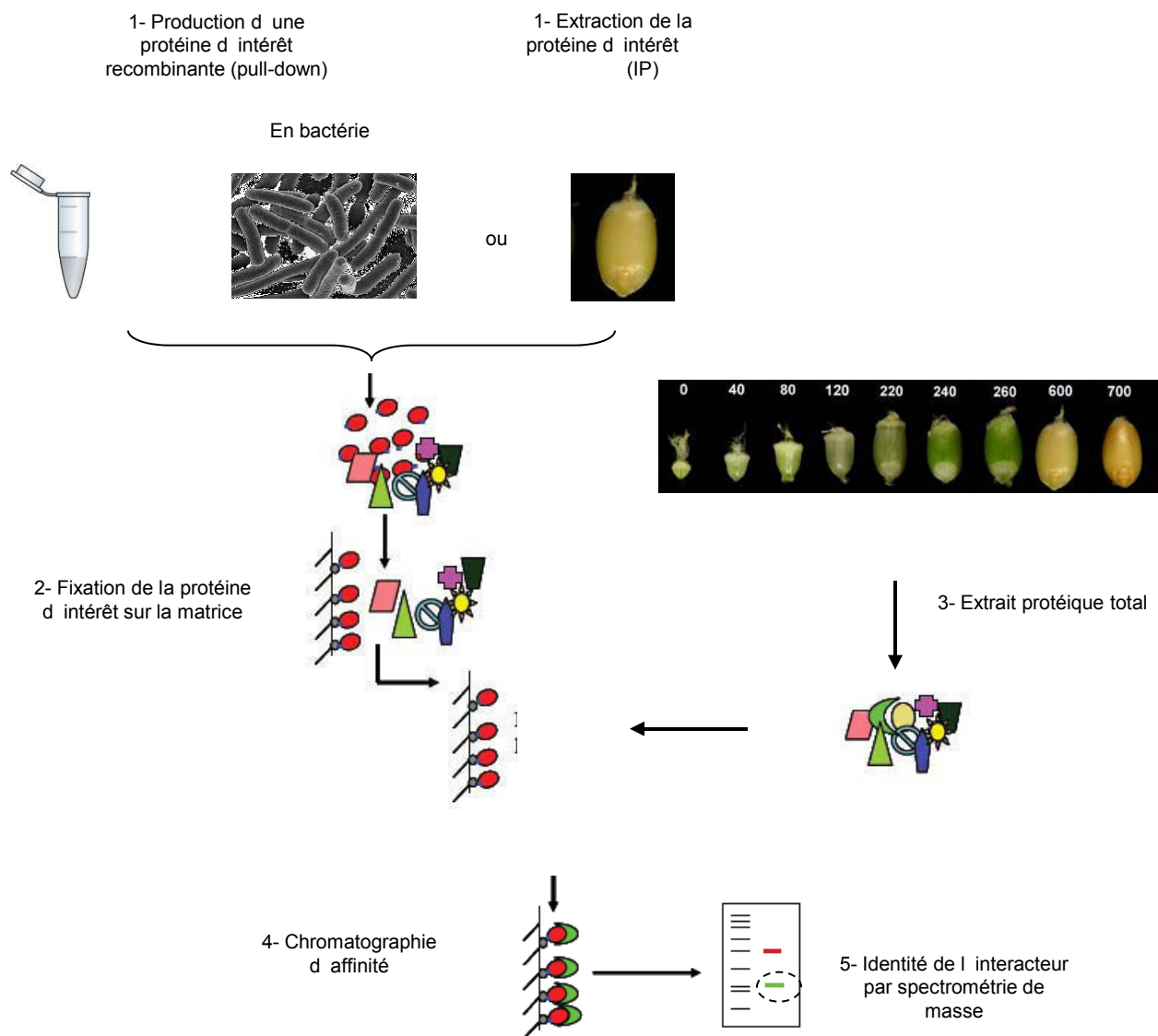


Figure 87 : Schéma simplifié du protocole expérimental de la recherche d'IPP par chromatographie d'affinité. La chromatographie d'affinité consiste en la fixation de la protéine d'intérêt à un support physique (ou matrice) par l'intermédiaire d'un anticorps dirigé contre la protéine (immunoprécipitation) ou contre un tag de la protéine-fusion d'intérêt (pull-down), puis au passage d'un extrait protéique total sur la matrice : les protéines interagissant avec la protéine d'intérêt fixée sont retenues puis éluées. L'identité des protéines interactrices éluées est finalement déterminée par spectrométrie de masse.

Chez le riz, Song et ses collaborateurs ont suggéré que la protéine GW2 était un régulateur négatif de la division cellulaire dans les enveloppes externes du grain. Les auteurs ont montré que GW2 est une E3 ubiquitine ligase fonctionnelle *in vitro* mais les substrats ou partenaires protéiques de GW2 restent à ce jour encore inconnus. Nos résultats de validation fonctionnelle de *TaGW2* par ARN interférence montrent que *TaGW2* est un régulateur positif de la taille et du poids du grain, et suggèrent que la sous-expression de *TaGW2* a un effet négatif sur le nombre de cellules de l'albumen du grain. Afin de déterminer la (ou les) voie(s) de signalisation ou processus cellulaire(s) dans lesquels la protéine TaGW2 intervient au cours du développement du grain, nous avons initié la recherche de ses interacteurs protéiques dans le grain de blé tendre.

Différentes stratégies peuvent être envisagées pour détecter des interactions protéine-protéine (IPP). Au cours de la dernière décennie, différentes approches biochimiques (TAP-TAG, immunoprécipitation, puces à protéines, librairie de peptides de séquence aléatoire) et moléculaires (double-hybride) ont été décrites pour la caractérisation de protéines interagissant avec une protéine d'intérêt (Puig *et al*, 2001 ; Gavin *et al*, 2002 ; Gupta *et al*, 2007 ; Persaud et Rotin, 2011 ; Guo *et al*, 2012). La recherche d'IPP peut être effectuée avec *a priori*, c'est-à-dire que l'existence d'une IPP est recherchée pour deux partenaires protéiques connus. Dans le cas de notre étude, cependant, aucune piste bibliographique ne permet pour l'instant de valider des interacteurs potentiels de TaGW2. La recherche des partenaires protéiques de TaGW2 nécessite donc une recherche d'IPP sans *a priori*.

Deux techniques majeures ne nécessitant pas l'acquisition d'équipements spécifiques coûteux et pouvant être rapidement mises en place peuvent permettre de répondre à notre problématique : la chromatographie d'affinité couplée à la spectrométrie de masse (ou AP-MS, Affinity Purification Mass Spectrometry) et le criblage de banque ADNc pleine-longueur par la technique de double-hybride (détaillée dans le Matériel et Méthodes, p.81). Nous avons montré précédemment dans le chapitre I que TaGW2 est une E3 ligase *in vitro*. Par conséquent, le substrat protéique ciblé et ubiquitiné par TaGW2 peut être potentiellement dégradé *in vivo* par l'E3 ligase. La détection d'IPP par chromatographie d'affinité se faisant à partir d'un extrait protéique total (Figure 87), la recherche d'interacteurs de TaGW2 au cours du développement du grain nécessiterait de devoir bloquer la voie ubiquitine-protéasome 26S idéalement uniquement dans le grain et de connaître précisément la « fenêtre » d'expression protéique du partenaire au cours du développement du grain. Cependant, nous ne pouvons actuellement pas nous aider du profil d'accumulation protéique de TaGW2 pour prédire la fenêtre précise d'expression des protéines interactrices potentielles de TaGW2, d'autant plus

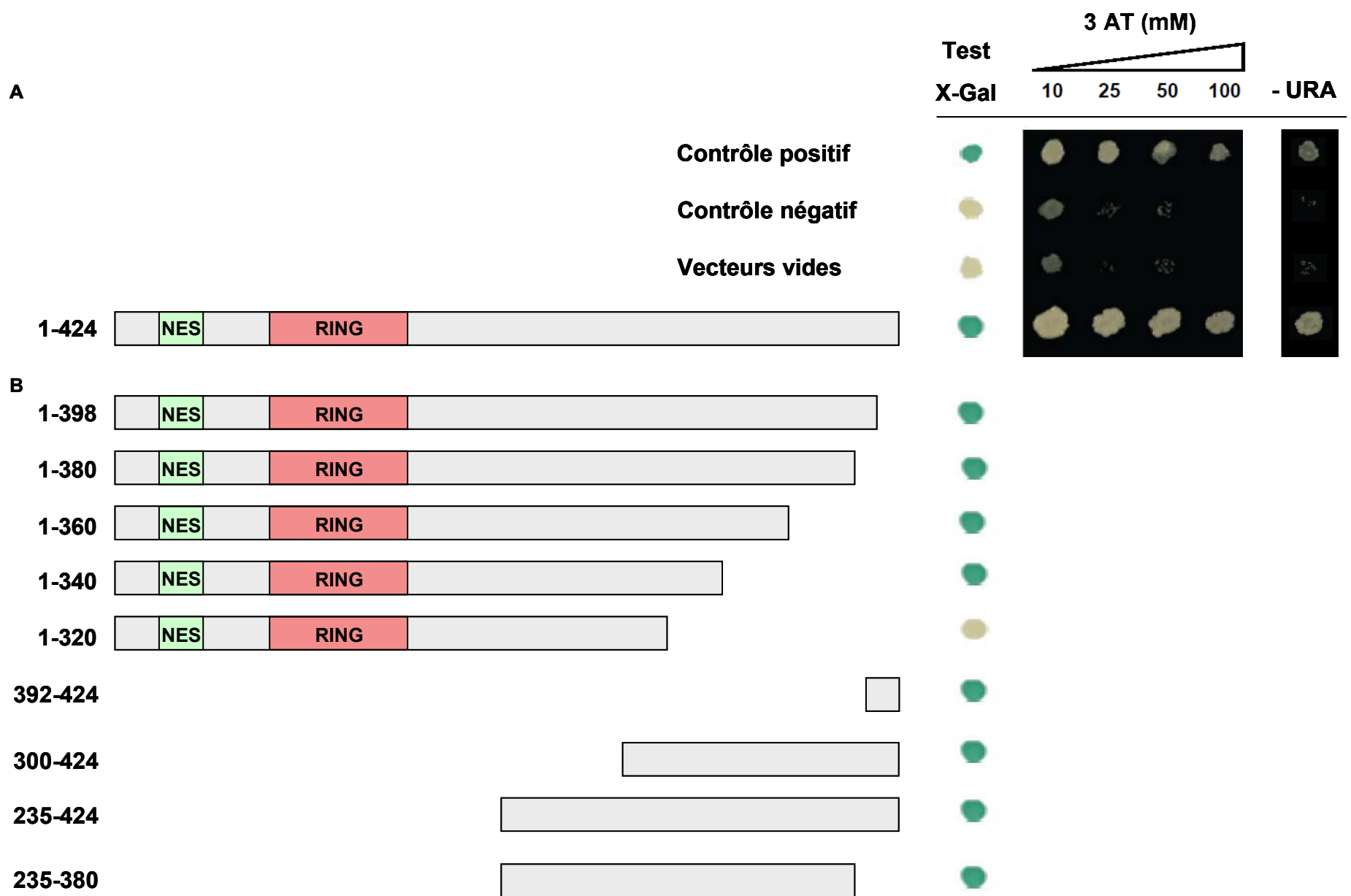


Figure 88 : Détermination de la séquence maximale non-autoactivatrice de TaGW2. A gauche, représentation schématique de la séquence entière de TaGW2 (A) et des mutants de délétion de TaGW2 (B) testés pour l'activation des gènes rapporteurs. L'activation du gène rapporteur *LacZ*, représentée par une coloration bleue de la colonie, a été testée sur un milieu –L/-T/X-Gal. L'activation des gènes rapporteurs nutritionnels *HIS* et *URA* ont été testés respectivement sur des milieux –L/-T/H contenant des concentrations croissantes de 3 AT, et un milieu –L/-T/H/-A. Contrôle positif : p32-Krev1 + p22-RalGDSwt ; contrôle négatif : p32-Krev1 + p22-RalGDS-m2.

si ces protéines, même si elles ne sont pas dégradées, ont une demi-vie très courte. De plus, l'administration d'inhibiteurs du protéasome uniquement dans le grain ne peut pas se faire sur plante entière et nécessite le développement de systèmes de culture *in vitro* d'organes comme la culture *in vitro* d'épis de blé. Dans notre contexte scientifique, la recherche d'interacteurs de TaGW2 par criblage de banque d'ADNc par la technique de double-hybride est plus aisée et plus rapide. Nous avons donc construits deux banques d'ADNc pleine longueur à partir d'ARN messagers de grains de blé aux stades de développement « division » et « remplissage », dans deux systèmes de criblage différents (double-hybrides nucléaire et cytoplasmique). Le criblage de la banque « division » dans les deux systèmes double-hybride avec TaGW2-A comme protéine appât a été initiée.

1. Détermination de la séquence de *TaGW2* à utiliser dans le système double-hybride nucléaire

L'un des pré-requis à l'utilisation d'un système double-hybride nucléaire pour la recherche d'IPP pour une protéine d'intérêt est de vérifier que la séquence protéique chimère testée (séquence d'intérêt fusionnée au domaine de liaison à l'ADN du facteur de transcription GAL4) n'active pas la transcription des gènes rapporteurs indépendamment d'une interaction avec une protéine proie chimère. En effet, il arrive que des protéines présentent des capacités d'« auto-activation », c'est-à-dire la capacité de recrutement de la machinerie de transcription au promoteur des gènes rapporteurs. Ceci peut être le cas des facteurs de transcription par exemple, mais également de protéines n'intervenant pas naturellement dans le processus de transcription mais pouvant le faire dans un système double-hybride classique du fait de domaines amphipathiques acides (Ruden *et al*, 1991 ; Ruden, 1992).

Nous avons testé cette propriété d'auto-activation pour TaGW2 dans le système double-hybride nucléaire Proquest (Invitrogen). La souche de levure MaV203 a été co-transformée avec le vecteur recombinant p32-TaGW2-A comportant la séquence complète codante de *TaGW2-A* clonée dans le vecteur pDEST32, en aval du domaine de liaison à l'ADN (DNA-BD) de GAL4, et le vecteur pDEST22 vide (comportant le domaine d'activation de GAL4). La séquence complète codante de *TaGW2-A* auto-active l'expression des 3 gènes rapporteurs et ne peut donc pas être utilisée pour la recherche d'IPP de TaGW2-A dans un système double-hybride classique (nucléaire ; Figure 88, A). Afin de rechercher la séquence codante maximale de *TaGW2-A* n'auto-activant pas dans un système double-hybride nucléaire, nous avons généré des mutants de délétions de *TaGW2-A* et les avons testé pour

l' « auto-activation » du gène *LacZ* (Figure 88, B). La séquence protéique de l'acide aminé 1 à 320 de TaGW2-A est la séquence codante maximale n'auto-activant pas le gène *LacZ*.

La recherche de partenaires protéiques de TaGW2-A dans un système double-hybride nucléaire implique donc d'utiliser une séquence partielle de *TaGW2-A*, dans laquelle il manque le dernier quart de la séquence protéique de *TaGW2-A* (acides aminés 321 à 424). Aucun domaine protéique conservé connu n'a pu être identifié dans cette région (AA321-424). Cependant, cette région pourrait contenir un domaine putatif de reconnaissance du substrat protéique ciblé par TaGW2-A. Afin d'optimiser nos chances d'identifier la ou les protéine(s) ciblée(s) par TaGW2-A par technique de double-hybride, nous avons alors choisi un système adapté à la recherche d'IPP de protéines qui sont fortement auto-activatrices en système classique : le double-hybride cytoplasmique ou « split-ubiquitine ». Ce système permet de pouvoir potentiellement utiliser la pleine-longueur de la séquence qui auto-active en double-hybride nucléaire. L'utilisation de ce système a également un autre avantage : les IPP ont lieu dans le cytoplasme. TaGW2 présentant à la fois une localisation subcellulaire dans le noyau et le cytoplasme, ce système permet de se rapprocher des conditions d'interaction de TaGW2 avec une protéine cible potentielle dans le cytoplasme.

2. Recherche des partenaires protéiques de TaGW2 par Y2H cytoplasmique

2.1. Construction de la banque ADNc normalisée

Une banque d'ADNc « division » a été construite à partir de l'ensemble des ARN messagers de grains de blé tendre entiers récoltés à 80, 160 et 220°Jours, dans un système de clonage permettant un criblage par double-hybride split-ubiquitine (système DUALhunter, Dualsystems Biotech). Afin d'optimiser les chances d'identifier un interacteur potentiel de TaGW2-A faiblement exprimé, cette banque a été normalisée. Le titrage de cette banque dénombre $9,2 \cdot 10^6$ clones indépendants, ce qui est largement au-dessus du minimum préconisé par le fournisseur ($2 \cdot 10^6$ clones indépendants). Une estimation de la taille moyenne des inserts de la banque par digestion enzymatique montre une taille moyenne de 2000 pb, supérieure à la taille moyenne des inserts (1200 pb) recommandée par le fournisseur. La banque d'ADNc construite a donc été estimée de bonne qualité.

Tableau 13 : Tests fonctionnels n° 1 et 2 de la construction Ost4-TaGW2A-Cub-LexA(VP16) dans le système Split-Ubiquitine DualHunter.

Ost4-TaGW2A-Cub-LexA(VP16)

Vecteur "proie"	% croissance sur			% croissance sur	
	SD/-L/-T	SD/-L/-T/-H	SD/-L/-T/-H	SD/-L/-T/-H/-A	SD/-L/-T/-H/-A
+ Ost1-Nubl (test 1)	3700	1600	43%	450	12%
+ pPR3N (test 2)	3500	0	0%	0	0%

2.2. Vérification de la faisabilité de recherche d'IPP de TaGW2 en double-hybride split-ubiquitine

La séquence complète codante de *TaGW2-A* a été clonée en partie C-terminale de la protéine d'ancrage à la membrane du réticulum endoplasmique Ost4, et en partie N-terminale de l'ensemble « moitié C-terminale de l'ubiquitine (Cub) – facteur de transcription LexA(VP16) », dans le vecteur pDHB1 (Dualsystems Biotech, Suisse). Afin de vérifier que la protéine chimère TaGW2-A est correctement exprimée dans la levure et qu'elle peut être utilisée dans un système split-ubiquitine, nous avons procédé à des tests fonctionnels dans la souche de levure NMY51 (Tableau 13). Pour cela, la souche NMY51 est co-transformée avec la construction Ost4-TaGW2A-Cub-LexA(VP16) et :

- 1) le contrôle Ost1-NubI (test n°1),
- 2) le plasmide vide de la banque ADNc, pPR3N, comportant NubG (test n°2).

NubI est la version sauvage de la partie N-terminale de l'ubiquitine, et possède naturellement une affinité très forte pour la partie C-terminale de l'ubiquitine Cub. NubG est une version mutante de NubI et ne peut pas reformer spontanément une ubiquitine avec Cub.

Le test n°1 permet de voir l'expression de la protéine chimère dans la souche de levure NMY51. Si la protéine chimère Ost4-TaGW2A-Cub-LexA(VP16) est correctement exprimée dans la levure, bien insérée dans la membrane du réticulum endoplasmique et que l'ensemble « Cub-LexA(VP16) » est correctement exposé sur la face cytoplasmique de la membrane du réticulum endoplasmique, les 2 moitiés NubI et Cub interagissent, reformant un monomère d'ubiquitine (split-ubiquitine). L'ubiquitine reformée est alors reconnue par une enzyme UBP (Ubiquitin specific protease), qui clive et libère le facteur de transcription LexA(VP16). Le facteur de transcription LexA(VP16) peut librement aller dans le noyau pour activer la transcription des gènes rapporteurs. L'expression correcte de Ost4-TaGW2A-Cub-LexA(VP16) a été testée sur les milieux –L/-T/-H et –L/-T/-H/-A. Une expression correcte de la protéine d'intérêt est estimée pour une croissance sur ces milieux sélectifs comprise entre 10 et 100%, d'après le fournisseur. Les levures co-transformées avec Ost4-TaGW2A-Cub-LexA(VP16) et le contrôle Ost1-NubI présentent un pourcentage de croissance de 43% sur milieu –L/-T/-H et de 12% sur milieu –L/-T/-H/-A (Tableau 13). Ost4-TaGW2A-Cub-LexA(VP16) semble donc peu exprimée dans la souche NMY51 ; ou la conformation tridimensionnelle de Ost4-TaGW2A-Cub-LexA(VP16) ne permet pas une bonne accessibilité au résidu Cub.

Tableau 14 : Test de détermination de la composition du milieu sélectif à utiliser (test n° 3) pour le criblage de la banque « division » avec la protéine appât TaGW2, dans le système double-hybride split-ubiquitine DualHunter.

Milieu de sélection	Nombre de colonies
SD/-L/-T/-H	30
SD/-L/-T/-H + 1mM 3AT	30
SD/-L/-T/-H + 2,5mM 3AT	5
SD/-L/-T/-H + 5mM 3AT	3
SD/-L/-T/-H + 7,5mM 3AT	10
SD/-L/-T/-H + 10mM 3AT	3
SD/-L/-T/-H/-A	1
SD/-L/-T/-H/-A + 1mM 3AT	0
SD/-L/-T/-H/-A + 2,5mM 3AT	0
SD/-L/-T/-H/-A + 5mM 3AT	0
SD/-L/-T/-H/-A + 7,5mM 3AT	0
SD/-L/-T/-H/-A + 10mM 3AT	0

Le pourcentage de croissance sur –L/-T/-H/-A étant tout de même dans l'intervalle donné par le fournisseur, nous avons décidé de procéder au 2^{ème} test et de tenter de cribler la banque ADNc construite.

Le test n°2 permet de vérifier que la construction Ost4-TaGW2A-Cub-LexA(VP16) n'auto-active pas les gènes rapporteurs (i.e qu'il n'y ait pas d'expression des gènes rapporteurs en absence d'IPP). En effet, certaines protéines peuvent posséder une forte capacité d'auto-activation dans ce système du fait de leur instabilité (dégradation), souvent causée par leur surexpression dans ce système, ou de la présence de séquences reconnues par l'enzyme UBP. Aucune colonie co-transformée n'a été détectée sur les milieux –L/-T/-H et –L/-T/-H/-A (Tableau 13). La protéine chimère Ost4-TaGW2A-Cub-LexA(VP16) n'auto-active donc pas les gènes rapporteurs HIS3 et ADE2 dans le système split-ubiquitine.

Avant de procéder au criblage de la banque ADNc, un 3^{ème} test a été réalisé afin de déterminer la composition du milieu sélectif le mieux adapté au criblage de la banque, c'est-à-dire celui présentant le moins de « bruit de fond » de croissance, et pour limiter la détection ultérieure de faux-positifs. Après sous-transformation de la souche NMY51 transformée pour Ost4-TaGW2A-Cub-LexA(VP16), avec le vecteur de la banque pPR3N vide, nous avons obtenu un nombre total de transformants de $1,62.10^7$ cfu (*colony units formed*, nombre de colonies formées), supérieur au nombre minimal de transformants préconisés par le fournisseur (8.10^5 cfu). Une seule colonie est apparue sur le milieu sélectif –L/-T/-H/-A puis aucune sur les milieux –L/-T/-H/-A supplémentés en 3AT (Tableau 14). Le milieu sélectif –L/-T/-H/-A + 1mM 3AT a par conséquent été choisi pour le criblage de la banque ADNc « division » par split-ubiquitine, avec la protéine TaGW2-A complète comme protéine appât.

2.3. Criblage de la banque ADNc « division » avec la protéine TaGW2-A pleine-longueur, par split-ubiquitine

La souche de levure NMY51 transformée pour Ost4-TaGW2A-Cub-LexA(VP16) a été sous-transformée avec la banque ADNc « division ». Le nombre total de transformants était de $4,8.10^6$ cfu et l'efficacité de transformation de $1,7.10^5$ clones/ μ g d'ADN, supérieurs aux minimums requis par le fournisseur (soit respectivement 2.10^6 cfu et $1,5.10^5$ clones/ μ g d'ADN). Étonnamment, aucune colonie n'a été obtenue sur les boîtes de milieu sélectif –L/-T/-H/-A + 1mM 3AT.

La tentative de criblage de la banque ADNc a donc été refaite en étalant sur des boîtes de milieu sélectif –L/-T/-H/-A, moins restrictif que précédemment. Malgré un nombre total

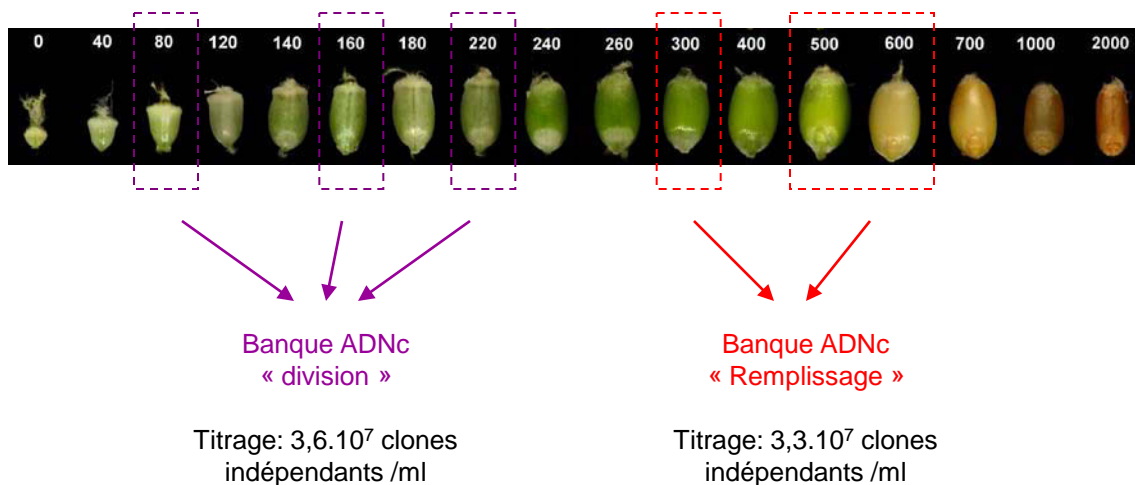






Figure 89 : Stades de développement de grains utilisés pour la création des deux banques ADNc pour un criblage double-hybride nucléaire par le système Matchmaker. Le titre de chacune des banques fabriquées est donné.

Tableau 15 : Détermination de la non-autoactivation des gènes rapporteurs MEL1 et AUR1C, de la protéine appât TaGW2(1-320).

<i>Constructions</i>	<i>Milieux de sélection</i>	
	-T/X	-T/X/Aur
Contrôle positif		
Contrôle négatif		X
DBD-TaGW2(1-320)		X

de transformants de $4,32.10^6$ cfu et une efficacité de transformation de $1,54.10^5$ clones/ μ g d'ADN, aucune colonie n'a été obtenue sur le milieu sélectif –L/-T/-H/-A.

2.4. Validation de la banque ADNc « division » construite dans le système DUALHunter

Afin de vérifier que l'échec d'obtention de clones « interacteurs » pour TaGW2-A ne soit pas dû à un problème dans la banque ADNc construite, nous avons criblé indépendamment cette banque ADNc « division » avec 5 autres protéines étudiées au laboratoire. Des colonies ont été obtenues sur les milieux sélectifs appropriés suite au criblage de la banque ADNc par chacune de ces protéines. La banque ADNc « division » construite peut donc bien être utilisée pour des criblages en split-ubiquitine avec des protéines appâts.

3. Recherche des partenaires protéiques de TaGW2 par Y2H nucléaire

3.1. Construction des banques ADNc

Deux banques d'ADNc ont été construites dans un système de clonage permettant un criblage par double-hybride nucléaire (système Matchmaker, Clontech ; Figure 89) :

- Une banque ADNc « division », à partir de l'ensemble des ARN messagers de grains de blé tendre entiers récoltés à 80, 160 et 220°CJours,
- Une banque ADNc « remplissage », à partir de l'ensemble des ARN messagers de grains de blé tendre entiers récoltés à 300, 500 et 600°CJours.

Ces banques sont « hébergées » dans des levures de souche Y187 ; et leur criblage est réalisé en exploitant la compatibilité sexuelle ou « mating » avec une souche Y2HGold transformée avec la protéine d'intérêt appât. Les titrages des banques ADNc « division » et « remplissage » dénombrent respectivement $3,6.10^7$ clones indépendants et $3,3.10^7$ clones indépendants /ml de suspension de levures de la banque (le minimum requis par le fournisseur est de 2.10^7 clones indépendants /ml). Une estimation de la taille moyenne des inserts de la banque par PCR montre une taille moyenne de 900 pb.

3.2. Vérification de la non-autoactivation de la protéine chimère DBD-TaGW2(1-320)

La séquence partielle de l'ADNc de *TaGW2-A*, des nucléotides 1 à 960 (acides aminés 1 à 320) a été clonée en partie C-terminale du domaine de liaison à l'ADN du facteur de

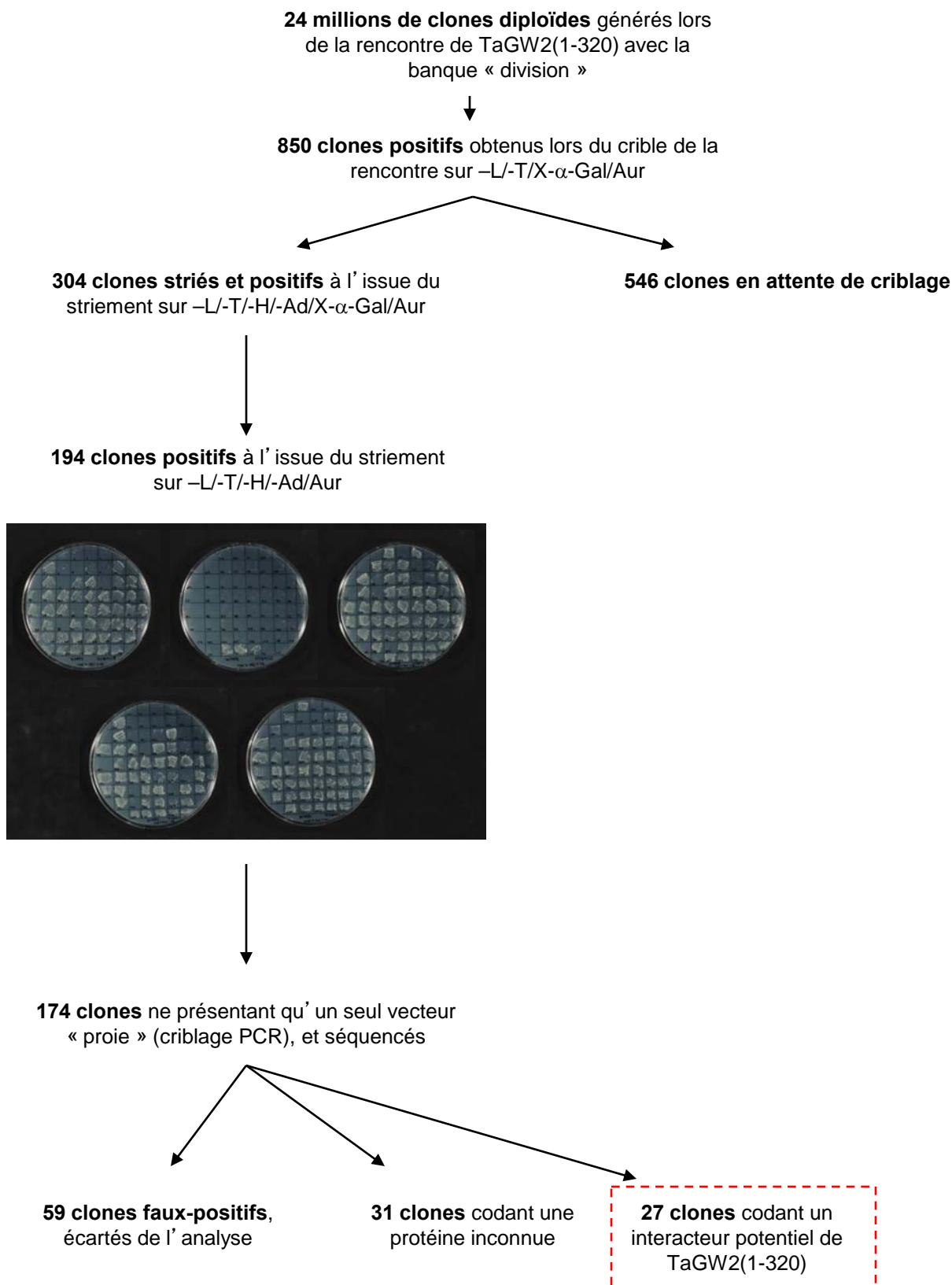


Figure 90 : Schéma récapitulatif du crible double-hybride de la banque « division » avec TaGW2(1-320) comme protéine appât.

transcription GAL4 (DNA-BD GAL4), dans le vecteur pGBKT7, puis la souche Y2HGold a été transformée avec le vecteur recombinant. La non-autoactivation des gènes rapporteurs (*MEL1* et *AURI-C*) par cette protéine tronquée de TaGW2-A a été vérifiée sur des milieux sélectifs (Tableau 15). Aucune croissance de la souche transformée n'a été observée sur le milieu –T/X/Aur. Cependant la souche de levure est bleu pâle sur le milieu –T/X. La protéine fusion DBD-TaGW2(1-320) n'auto-active donc pas le gène rapporteur *AURI-C* mais auto-active faiblement le gène *MEL1*. Le criblage de la banque se faisant sur un milieu sélectif contenant toujours de l'Auréobasidine A, la faible auto-activation de la protéine fusion pour le gène *MEL1* ne gêne pas pour un criblage de banque avec DBD-TaGW2(1-320) comme protéine appât.

3.3. Criblage de la banque « division » avec TaGW2-A(1-320)

Le criblage de la banque « division » avec la protéine TaGW2-A(1-320) comme protéine appât a été réalisée par « mating ». La souche haploïde Y2HGold, de type sexuel MATa, transformée avec la construction DBD-TaGW2-A(1-320), est mise au contact de la souche haploïde Y187, de type sexuel MAT α , renfermant la banque « division ». Le titrage de cette rencontre, c'est-à-dire le nombre total de diploïdes formés au cours de cette rencontre, a été estimé sur boîtes de milieu –L/-T et était de $2,4.10^7$ cellules diploïdes, nombre supérieur au minimum requis par le fournisseur (1.10^6). Le criblage de ces diploïdes a été entrepris sur le milieu sélectif –L/-T/X- α -Gal/Aur. Ce milieu sélectionne les cellules diploïdes exprimant les gènes rapporteurs *MEL1* et *AURIC*. Il permet donc une sélection uniquement des cellules diploïdes comportant des protéines interactrices potentielles de TaGW2-A(1-320), puisque l'expression des gènes rapporteurs n'est permise que lorsqu'il y a en théorie rapprochement dans l'espace des domaines DBD et AD de Gal4 du fait de l'interaction de la protéine appât avec une protéine proie de la banque.

Près de 850 clones indépendants ont été dénombrés sur les boîtes de milieu sélectif –L/-T/X- α -Gal/Aur, à l'issue de la rencontre (Figure 90). Devant le très grand nombre de clones indépendants obtenus, nous avons dans un premier temps choisi de ne traiter que 304 clones. Tous les clones se sont développés après avoir été striés sur milieu –L/-T/-H/-Ad/X- α -Gal/Aur. Un striement de l'ensemble de ces clones sur milieu –L/-T/-H/-Ad/ Aur a permis d'éliminer des clones et de n'en conserver que 194. Une PCR sur colonies sur chacun des 194 clones, afin de distinguer le nombre de vecteurs « proie » contenus dans chaque clone, a

Tableau 16 : Résultat de l'analyse blastx après séquençage des 174 clones interacteurs potentiels de TaGW2(1-320).

Catégories de protéines criblées identifiées	Nombre de clones
Sous-unité beta de l'ATP synthase **	26
Cytochrome P450 **	23
Sous-unité d'E3 ligase (sous-unité adaptatrice d'un complexe CUL3-BTB)	1
Facteurs de transcription	2
<i>de type NAC</i>	1
<i>de type MYB (famille R1R2R3)</i>	1
Histones **	2
Métabolisme secondaire	1
<i>betaine-aldehyde dehydrogenase</i>	
Protéine de la division cellulaire	2
<i>sous-unité p60 ATPase de la katanine</i>	1
<i>homologue de la protéine fstZ</i>	1
Protéine impliquée dans le contrôle de la vernalisation	1
<i>VIN3-like protein (VIL3)</i>	
Signalisation auxine	1
<i>Flavin monooxygenase YUCCA11</i>	
Ubiquitine ou Ubiquitine-like	2
Protéines métaboliques du grain de blé	2
Protéines diverses dont:	35
<i>Actine</i>	1
<i>Protéine du métabolisme lipidique</i>	1
<i>Protéines impliquées dans le transfert des lipides *</i>	7
<i>Protéine de la traduction</i>	1
<i>Enzymes régulant la conformation de protéines</i>	1
<i>Protéines chloroplastiques</i>	5
<i>Sous-unité ribosomale *</i>	1
<i>Protéines mitochondriales **</i>	3
<i>Protéine du métabolisme carboné (GAPDH)</i>	1
<i>Protéines de défense</i>	10
<i>Protéases **</i>	2
Protéines de réserves du grain de blé * dont:	11
<i>Gliadines *</i>	8
<i>Gluténines de faible poids moléculaire *</i>	3
Protéines inconnues	15

*, faux-positifs révélés suite au test 2 à 2 d'un crible ultérieur (i.e activation des gènes rapporteurs par le clone proie indépendamment de la présence de la protéine appât).

**, faux-positifs probables, i.e interacteur positif d'une protéine appât différente de TaGW2(1-320), révélé suite au test 2 à 2 d'un crible ultérieur (i.e activation des gènes rapporteurs uniquement en présence du clone proie et de la protéine appât).

permis d'écarter de nouveau des clones et 174 clones ne présentant qu'un seul vecteur « proie » ont été séquencés afin de connaître la nature de la séquence « proie ».

3.4. Analyse des séquences des clones interacteurs potentiels de TaGW2(1-320).

Une analyse blastx dans NCBI a été réalisée pour les 174 séquences des clones interacteurs potentiels de la protéine appât. Le résultat de cette analyse est présenté dans le tableau 16. Cinquante-neuf clones faux-positifs au total ont été détectés et ont été écartés pour la suite de l'analyse: 19 clones ayant été révélés faux-positifs lors du test 2 à 2 d'un précédent crible de la banque division avec une autre protéine appât que TaGW2 et 40 clones ayant été révélés interacteurs positifs d'une protéine appât différente de TaGW2(1-320) à l'issue d'un test 2 à 2 d'un crible ultérieur. Les clones de cette dernière classe étant commun à plusieurs protéines appât totalement différentes, ils ont été considérés comme des faux-positifs hautement probables. Ce type de protéines sont souvent des protéines présentant une affinité naturelle non-spécifique pour d'autres (ou dites « collantes »), ou peuvent être des protéines qui, n'étant pas dans leur compartiment subcellulaire particulier d'origine, adoptent une conformation qui les rendent « collantes ». L'analyse blastx a permis de révéler 31 clones codant pour des protéines inconnues et 27 clones codant des protéines présentant des homologies de séquences avec des protéines connues. Parmi ces 27 clones, on peut noter notamment la présence de :

- Une sous-unité adaptatrice (BTB) d'un complexe E3 CUL3-BTB
- Une ubiquitine et une ubiquitine-like
- Deux facteurs de transcription, de type NAC et MYB
- Une enzyme du métabolisme secondaire (synthèse de bétanine)
- Un homologue d'une enzyme de la voie de synthèse de l'auxine (YUCCA11)
- Deux protéines de la division cellulaire : une protéine homologue à une sous-unité ATPasique de la katanine chez les mammifères, et une protéine homologue de la protéine ftsZ chez les bactéries.

Conclusions et discussions du Chapitre III.

Afin de rechercher les partenaires protéiques de TaGW2 au cours du développement du grain, nous avons généré trois banques d'ADNc à partir de grains de blé tendre à différents stades de développement, pour leur criblage dans deux systèmes de double-hybride : une banque « division » et une banque « remplissage » générées dans un système pour un criblage par double-hybride nucléaire ; et une banque « division » générée dans un système pour un criblage par double-hybride cytoplasmique. La nature de la protéine TaGW2 ne permettant pas un criblage par double-hybride nucléaire avec la totalité de la séquence protéique de TaGW2, nous nous sommes dans un premier temps concentrés sur un criblage de la banque « division » par double-hybride cytoplasmique. Malheureusement, cette tentative a échoué à plusieurs reprises, peut être du fait de la toxicité de la protéine complète TaGW2-A dans la souche de levure NMY51 ou d'une mauvaise accessibilité de la protéine appât chimère au système de criblage. La banque « division » a donc été criblée par double-hybride nucléaire avec les trois-quarts de la protéine TaGW2-A comme protéine appât. Les premiers résultats ont permis d'identifier 27 clones codant des protéines interactrices potentielles de TaGW2.

Cette liste n'est qu'une première étape dans le crible double-hybride de la banque ADNc. En effet, cette liste contient très probablement de nombreux faux-positifs qui doivent être éliminés par un test 2 à 2. Dans ce test 2 à 2, pour une protéine proie considérée, des levures Y2HGold sont co-transformées avec :

- 1) la protéine proie potentielle et la protéine appât DBD-TaGW2(1-320)
- 2) la protéine proie potentielle et le vecteur DBD vide.

Si pour une protéine proie considérée, les levures transformées poussent quelque soit la présence ou l'absence de la protéine appât DBD-TaGW2(1-320) sur les milieux de sélection pour les gènes rapporteurs ; alors cette protéine proie est un faux-positif. Les protéines codées par les clones restants à l'issue de ce test 2 à 2 peuvent alors être considérées comme des interacteurs positifs potentiels de la protéine appât. Cependant, cette interaction reste potentielle et nécessite toujours d'être confirmée par une autre technique d'analyse d'interaction protéine-protéine, comme par exemple le pull-down ou encore le BiFC. En effet, les clones révélés comme interacteurs positifs à l'issue du test 2 à 2 peuvent coder des protéines « collantes », qui présentent une affinité aspécifique pour un grand nombre de protéines.

Un facteur de transcription, cible de TaGW2 ?

Deux séquences de facteurs de transcription, un de type NAC et l'autre de type R1R2R3-MYB, ont été retrouvées à l'issue du crible. Les E3 ligases ont été montrées de longue date comme régulant de nombreux facteurs de transcription (pour exemple, Osterlund *et al*, 2000 ; Qin *et al*, 2008 ; Liu et Stone, 2010 ; Santner et Estelle, 2010), impliqués dans des processus de développement très différents ; et il n'est donc pas étonnant qu'une des cibles de TaGW2 puisse être un facteur de transcription. Malheureusement, les résultats des analyses blastx dans NCBI de ces deux séquences de facteurs de transcription ne reportent d'identité de séquences avec aucun autre facteur de transcription NAC ou R1R2R3-MYB déjà fonctionnellement caractérisé chez d'autres espèces. Il est donc relativement difficile pour l'instant de formuler des hypothèses quant à l'une des voies de signalisation dans laquelle pourrait intervenir TaGW2.

Les facteurs de transcription NAC (*NAM*, *ATAF1,2* et *CUC2*) sont spécifiques des plantes et constituent une des plus grandes familles de facteur de transcription (FT). Les génomes d'*Arabidopsis* et du riz comptent respectivement 117 et 151 gènes codant des FT NAC. Les FT NAC sont associés à de nombreux processus divers tels que plusieurs programmes de développement (entre autre embryogénèse, développement floral, développement racinaire) impliquant plusieurs voies de signalisation hormonale, la sénescence, la formation de parois secondaires, et les réponses aux stress biotiques et abiotiques (revues de Olsen *et al*, 2005 ; Puranik *et al*, 2012). Ces deux derniers semblent par ailleurs être les processus majoritairement régulés par les FT NAC (Puranik *et al*, 2012). Les FT NAC peuvent être régulés à différents niveaux (transcriptionnel, post-transcriptionnel et post-traductionnel), dont leur dégradation médiée par la voie UPS (revue de Puranik *et al*, 2012). Dans le développement de racines latérales en réponse à l'auxine, par exemple, la protéine E3 RING SINAT5 atténue le signal auxine, et donc la formation de racines latérales, en ciblant et polyubiquitinant le facteur NAC1 qui est ensuite dégradé par le protéasome 26S, réinitialisant ainsi la cascade de transduction (Xie *et al*, 2002).

De plus, les FT ANAC019 et ANAC055 (abscissic acid-responsive NAC 19 et 55) ont été montrés chez *Arabidopsis* interagissant avec la E3 ligase de type RING-H2, RHA2 (Greve *et al*, 2003 ; Bu *et al*, 2009). Il est encore à ce jour inconnu si les FT ANAC019 et ANAC055 sont ubiquitinés par RHA2. De manière intéressante, Greve et ses collaborateurs (2003) ont montré que le domaine RING de RHA2 était nécessaire et suffisant pour l'interaction avec ANAC019. Rappelons que notre crible double-hybride de la banque ADNc a été réalisé avec

la protéine TaGW2 incomplète, incluant tout de même le domaine RING. Dans l'hypothèse bien sûr que l'interaction entre le FT putatif NAC et TaGW2 soit confirmée, l'interaction pourrait peut-être se faire par l'intermédiaire du domaine RING de TaGW2.

L'import nucléaire des FT est un processus extrêmement régulé, et peut reposer entre autre sur la séquestration cytoplasmique des FT par leur ancrage à la membrane. Sous l'influence d'un signal, la levée de séquestration est alors opérée par clivage protéolytique du précurseur signal d'ancrage membranaire qui autorise alors la localisation nucléaire des FT. A ce jour, deux mécanismes de clivage du précurseur ont été reportés, dont un mécanisme dépendant de l'UPS, le RUP (Regulated Ubiquitin/proteasome-dependent Processing) (Hoppe et al, 2001 ; revue de Olsen *et al*, 2005). Une analyse des séquences primaires des NAC chez *Arabidopsis* et le riz, a révélé qu'un certain nombre de FT NAC possédait une région transmembranaire, indiquant que la rétention membranaire peut être un des moyens de régulation de l'import nucléaire des FT NAC (revue de Olsen *et al*, 2005). Il peut être intéressant de rappeler que la protéine TaGW2 présente une localisation nucléocytoplasmique ; et que le domaine RING semble important pour la localisation cytoplasmique de TaGW2, suggérant une activité E3 ligase de TaGW2 dans le cytoplasme. Dans l'hypothèse toujours que l'interaction entre le putatif FT NAC et TaGW2 soit confirmée, l'hypothèse suivante pourrait être formulée. TaGW2 pourrait participer au clivage du précurseur de ce FT NAC dans le cytoplasme, autorisant ainsi l'import nucléaire du FT NAC et donc son action sur la transcription de gènes. La localisation nucléaire de TaGW2 pourrait alors constituer un moyen de retenir TaGW2 pour l'empêcher d'agir en tant que E3 ligase dans le cytoplasme.

Les FT de type MYB constituent également une large famille de FT chez les plantes, avec par exemple 218 et 180 membres chez le riz et *Brachypodium*, respectivement (Yanhui *et al*, 2006 ; International Brachypodium initiative, 2010). Ils sont caractérisés par la présence d'1 à 4 répétitions d'un domaine caractéristique MYB (50 à 53 acides aminés dont 3 résidus tryptophane ou phénylalanine constituant une poche hydrophobe) en partie N-terminale de leur séquence primaire, et peuvent ainsi être classés en 4 catégories : les 4R-MYB avec 4 répétitions MYB, les 3R-MYB (R1R2R3-MYB) avec 3 répétitions MYB, et les R2R3-MYB et les MYB-related avec 2 et 1 répétition(s), respectivement (revue de Dubos *et al*, 2010). Les FT MYB sont impliqués dans de très nombreux processus physiologiques et biochimiques tels que la régulation du métabolisme primaire et secondaire, la réponse à des stress biotiques et

abiotiques, la synthèse d'hormone, la transduction de signaux, mais également le contrôle du cycle cellulaire (Stracke *et al*, 2001 ; Du *et al*, 2009 ; Dubos *et al*, 2010 ; Feller *et al*, 2011).

La séquence blé putative MYB isolée lors du crible double-hybride montre une homologie de séquence avec un membre R1R2R3-MYB. Plusieurs études ont montré que la famille R1R2R3-MYB, en particulier, était impliquée dans le contrôle du cycle cellulaire chez *Arabidopsis* et le tabac, au niveau de la transition G2/M (Ito *et al*, 2001 ; Ito, 2005 ; Haga *et al*, 2007, 2011). La régulation du cycle cellulaire est assurée par des points de contrôle impliquant des cyclines et des cyclines-dépendantes kinases (CDK ; revue de Inzé et De Veylder, 2006). Parmi ces protéines, certaines comme les cyclines des classes CYCB1, CYCB2 et CYCA1, ainsi que la protéine plante-spécifique CDKB2 sont exprimées spécifiquement à la transition G2/M du cycle cellulaire. Les gènes G2/M-spécifiques possèdent dans leur région promotrice des éléments *cis* communs appelés éléments activateurs mitose-spécifique, ou MSA (Mitosis-Specific Activator ; Ito *et al*, 1998 ; Ito, 2000). Chez le tabac, il a été montré que les FT de type R1R2R3-MYB, NtmybA1, NtmybA2 et NtmybB, pouvaient physiquement se lier à ces MSA (Ito *et al*, 2001). Le génome d'*Arabidopsis* compte 5 gènes codant des R1R2R3-MYB et la double mutation *myb3r1* et *myb3r4* résulte en la réduction du taux des transcripts des gènes spécifiques de la transition G2/M *CYCB1;2*, *CYCB2;1* et *CDC20.1*. Les double-mutants *myb3r1 myb3r4* présentent également de manière générale une cytokinèse incomplète dans la division somatique (Haga *et al*, 2007). Il a été montré que MYB3R1 et MYB3R4 régulaient la cytokinèse entre autre par l'activation de la transcription du gène *KNOLLE*, une syntaxin spécifique de la cytokinèse (Haga *et al*, 2007). Une analyse transcriptomique haut-débit du double-mutant *myb3r1 myb3r4* a révélé que de nombreux gènes autres que *KN*, impliqués et non-impliqués dans le contrôle du cycle cellulaire, étaient en réalité régulés par ces deux FT R1R2R3-MYB (Haga *et al*, 2011).

Une analyse dans les bases de données EST de blé afin de recenser toutes les séquences MYB a recensé seulement 72 séquences MYB, dont 1 seule de type R1R2R3-MYB, appelée TaMYB35. Zhang et ses collaborateurs (2012) ont montré que ce FT de type R1R2R3-MYB TaMYB35 était exprimé de manière ubiquitaire, et que son expression pouvait être modulée par un stress au froid. De manière intéressante, une analyse phylogénétique des facteurs MYB chez *Arabidopsis*, le riz et le blé tendre, a montré que TaMYB35 était phylogénétiquement très proche de AtMYB3R1 et 4, fonctionnellement caractérisés chez *Arabidopsis* dans le contrôle du cycle cellulaire.

La séquence que nous avons isolée à l'issue du crible double-hybride pourrait être TaMYB35. Dans l'hypothèse bien sûr que l'interaction entre ce putatif FT R1R2R3-MYB et

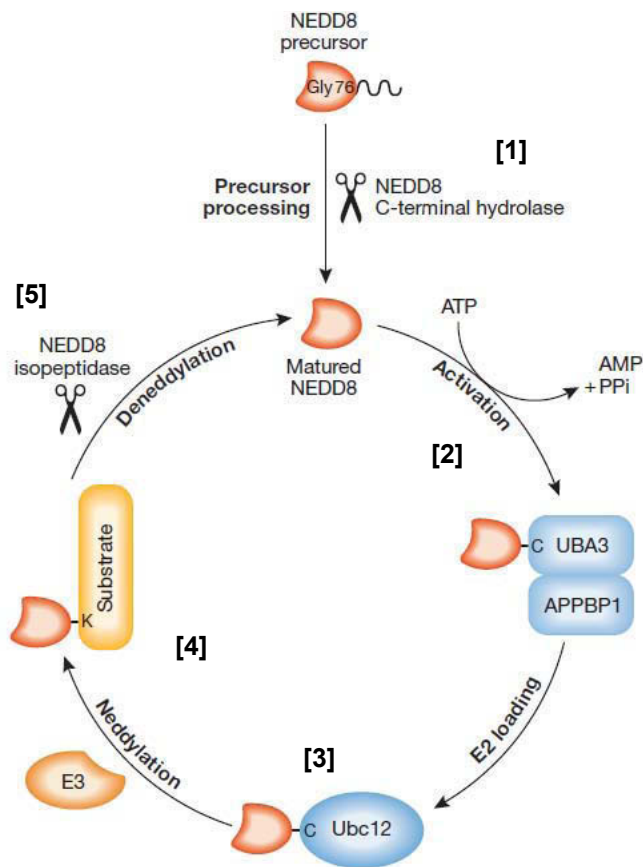


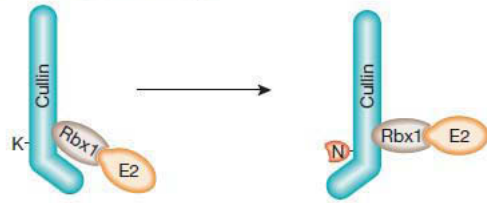
Figure 91 : Représentation schématique des principales étapes de la réaction de neddylation/deneddylation (d'après Rabut et Peter, 2008). La protéine NEDD8 mature est générée à partir du clivage d'un précurseur NEDD8 par une NEDD8 C-terminale hydrolase [1], puis activée par une E1 (l'hétérodimère UBA3/APPBP1 chez l'humain) selon une réaction dépendante de l'ATP [2]. La NEDD8 activée est alors transférée sur une E2 (Ubc12 chez l'humain) [3]. Une E3 NEDD8 ligase reconnaît la protéine substrat et appose le monomère NEDD8 sur celle-ci [4]. Le monomère NEDD8 peut être ensuite recyclé grâce à une NEDD8 isopeptidase [5].

TaGW2 soit confirmée, il peut être imaginé que TaGW2 impacte la taille du grain, par un contrôle de la division cellulaire dans l'albumen ou les téguments du grain, à travers ce FT R1R2R3-MYB. Il peut également être envisagé qu'à travers ce FT MYB, TaGW2 participe également à des processus de tolérance à des stress abiotiques. En effet, des FT R1R2R3-MYB (les homologues MYB3R1 et MYB3R4 chez *Arabidopsis* par exemple) ont été montrés dans la régulation de plusieurs gènes impliqués dans plusieurs processus.

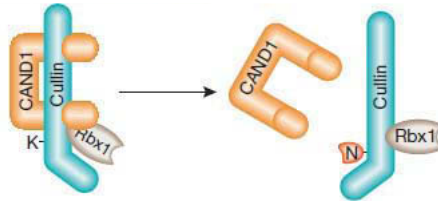
TaGW2, à la fois une E3 ubiquitine ligase et une E3 Nedd8 ligase ?

L'identification de l'ubiquitine au sein des différents clones interacteurs potentiels de TaGW2 n'est pas surprenante et tend même à confirmer que TaGW2 soit une E3 ligase. L'identification de la séquence de la protéine Nedd8, encore appelée, chez les végétaux, RUB1 (Related-UBiquitin1), dans les clones des interacteurs potentiels de TaGW2 est elle plus innovante quant à la fonction de TaGW2. Pour rappel, la protéine RUB1/Nedd8 est une protéine proche de l'ubiquitine de par sa conformation ; et la neddylation de protéines, c'est-à-dire l'apposition de monomère de RUB1/Nedd8 sur un substrat protéique, s'effectue selon un processus semblable à l'ubiquitination (Figure 91). Chez les plantes, la neddylation a été surtout largement rapportée dans la modification post-traductionnelle de la sous-unité culline des complexes CRL, contribuant à l'activation de ces complexes E3 ligases. Chez les animaux, de plus en plus d'études montrent que les cullines ne sont pas la seule classe de protéines à être modifiées par Nedd8 (revue dans Rabut et Peter, 2008). Tout comme l'ubiquitination, la neddylation a été montrée importante pour de nombreux processus cellulaires dont le contrôle de la division cellulaire chez les animaux et les plantes (Del Pozo *et al*, 1998, 2002 ; Tateishi *et al*, 2001). Rabut et Peter (2008) ont recensé les diverses fonctions de la neddylation de protéines animales (Figure 92). Les processus d'ubiquitination et de neddylation sont intimement liés. Par exemple chez les animaux, plusieurs exemples de protéines pouvant être à la fois ubiquitinées et neddyées ont été rapportés (p53, EGFR). De plus, il a été montré que des enzymes E3 pouvaient à la fois être des Ub et Nedd8 ligases fonctionnelles. La E3 ligase de type RING MDM2, par exemple, peut à la fois ubiquitiner et neddyler son substrat protéique p53 (Xirodimas *et al*, 2004). De manière identique, la E3 ligase de type RING c-CBL peut ubiquitiner et neddyler EGFR (Oved *et al*, 2006). De même, il est donc tout à fait envisageable que TaGW2 puisse être à la fois une E3 ubiquitine et NEDD8/RUB1 ligase.

A - Changements conformationnels



B - Blocage d'une autre interaction protéine-protéine



C - Création d'une nouvelle surface de liaison

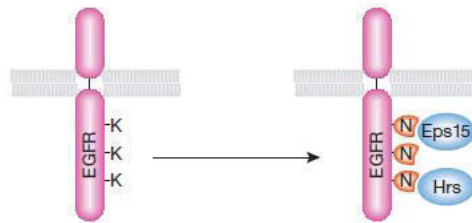


Figure 92: Les divers effets de la neddylation (d'après Rabut et Peter, 2008). La neddylation (N) peut induire des changements de conformation de ces cibles (A), permettant par exemple, dans le cas de la neddylation des cullines des complexes CRL animaux et végétaux, de positionner correctement l'enzyme E2 pour une polyubiquitination efficace du substrat protéique. De plus, la neddylation peut empêcher des interactions de la protéine modifiée avec d'autres partenaires (ex. CAND1, protéine inhibitrice de l'assemblage des complexes CRL) (B), ou encore entrer en compétition avec l'apposition d'autres modifications post-traductionnelles sur la protéine neddylée. Pour finir, la neddylation de protéines peut constituer sur la protéine modifiée des nouveaux sites de liaisons avec des interacteurs protéiques (C).

Tout comme des E3 ub ligases ont été montrées pouvant se réguler elles-mêmes par auto-ubiquitination, les E3 NEDD8 ligases peuvent également s'auto-neddylater (Xirodimas *et al*, 2004). Cette propriété de TaGW2 pourrait également expliquer la détection des clones codant l'ubiquitine et RUB1 dans le crible double-hybride.

Comme nous l'avons mentionné précédemment, le nucléole possède diverses fonctions dont la plus connue est la biogénèse des ribosomes. Les ARN ribosomaux sont transcrits dans le nucléole ; et les protéines ribosomales sont importées du cytoplasme vers le nucléole, où elles sont assemblées aux ARN ribosomaux pour former des ribosomes matures. De manière intéressante, Xirodimas et ses collaborateurs (2008) ont montré que de nombreuses protéines ribosomales étaient neddylées, et qu'il existait une relation très forte entre neddylation et ubiquitination des protéines ribosomales. Les auteurs suggèrent d'ailleurs que la neddylation contribuerait à la stabilisation des protéines ribosomales. De même, la localisation nucléolaire de TaGW2 pourrait être mise en lien avec sa possible double fonction Ub/Nedd8 ligase.

TaGW2-D dans la synthèse d'osmoprotectant au cours de la phase de maturation du grain ?

L'enzyme bétaine aldéhyde déshydrogénase (BADH) participe à la voie de synthèse de la glycine bétaine, en catalysant l'oxydation de la bétaine aldéhyde. La glycine bétaine (GB) fait partie de la famille des solutés compatibles. Ces solutés sont des osmo-protecteurs : ils s'accumulent dans la cellule, équilibrent les différences de potentiel osmotique entre l'intérieur et l'extérieur de la cellule, et protègent ainsi les protéines et les membranes cellulaires des dégâts engendrés par la dessiccation. La GB a été très étudiée chez les plantes dans la tolérance à des stress abiotiques (revue de Chen et Murata, 2010). Le compartiment d'accumulation de la GB n'est pour l'instant pas très bien établi et semble varier entre les espèces végétales. Il en est de même quant au compartiment de synthèse de ce soluté, mais il apparaîtrait que dans la majorité des cas les BADH présentent une accumulation chloroplastique ou cytoplasmique.

Lors de notre analyse transcriptomique, nous avons observé que le transcrit *TaGW2-D* continuait de s'accumuler à 800°CJours tandis que l'expression des copies A et B de *TaGW2* s'effondrait, suggérant un rôle plus particulier de *TaGW2-D* dans la maturation du grain. Cette forte accumulation coïncide notamment avec le pic d'accumulation de GB dans le grain de

blé, variété Hereward, décrit dans l'étude de Shewry (2012). Si l'interaction entre l'enzyme BADH et TaGW2 est confirmée, il pourrait être imaginé que TaGW2-D impacte la synthèse de GB en ubiquitinant, ou même neddylant, l'enzyme BADH.

TaGW2, une protéine E3 RING « multifonction » ?

Le très grand nombre de clones d'interacteurs potentiels de TaGW2(1-320) présentés en Tableau 16 va très probablement diminuer à la suite du test double-hybride 2 à 2. Dans l'hypothèse que les protéines mentionnées précédemment, voire peut-être même d'autres protéines présentées dans le tableau, soient confirmées comme interacteurs véritables de TaGW2 suite à des expériences ultérieures de pull-down ou de BiFC, TaGW2 pourraient donc interagir avec plusieurs protéines, impliquées dans des processus divers. Il peut d'ailleurs être souligné que ceci peut être cohérent avec une expression ubiquitaire du gène *TaGW2* dans la plante de blé tendre. Ces protéines interactrices pourraient être des protéines cibles de TaGW2, mais également des composants d'un complexe incluant TaGW2. Ce cas de figure d'une E3 RING « multifonction » n'est pas très rare et a déjà été relaté dans la littérature, comme dans le cas extrême par exemple de la protéine humaine BRCA1. La protéine BRCA1, incluse dans différents complexes, intervient dans de nombreux processus, interagissant avec de nombreuses protéines (Wu *et al*, 2008). La fonction E3 ligase de la protéine BRCA1 est assurée en hétérodimère avec la protéine BARD1, et l'hétérodimère est capable de produire de nombreux types d'ubiquitination en fonction de l'enzyme E2 interagissant avec BRCA1 (mono-, poly- en K48, K63, K6 ; Christensen *et al*, 2007 ; Wu *et al*, 2008). Ainsi, une situation similaire peut être imaginée pour TaGW2 qui interagirait avec une multitude de protéines, intervenant ainsi dans plusieurs processus en fonction des combinaisons d'interacteurs formées.

DISCUSSION GENERALE
ET
PERSPECTIVES

En 2007, le gène *GW2* a été fonctionnellement caractérisé chez le riz dans la régulation négative de la taille et du poids du grain de riz (Song *et al.*, 2007). Au moment de l'initiation de ce travail de thèse, aucune information concernant la présence et la fonction d'orthologues de *GW2* chez d'autres espèces n'était disponible et mon travail de thèse a donc consisté à rechercher la fonction de *TaGW2* chez le blé tendre, et plus particulièrement au cours du développement du grain chez cette espèce. Entre-temps, un des deux homologues de *GW2* chez le maïs, *ZmGW2-CHR4*, a été montré génétiquement lié au poids de 100-grains chez le maïs (Li *et al.*, 2010a) ; et deux études récentes chez le blé tendre ont rapporté l'implication de la copie A de *TaGW2* dans le contrôle du poids de 1000-grains chez le blé (Su *et al.*, 2011 ; Yang *et al.*, 2012).

Ce travail de thèse a permis de montrer tout d'abord que le gène *TaGW2* possédait trois copies génomiques portées par les trois génomes homéologues du blé tendre, sur les chromosomes 6AS, 6BS et 6DS. *TaGW2* est exprimé de manière ubiquitaire, et les trois copies *TaGW2-A*, *-B* et *-D* sont exprimées au cours du développement du grain. Cependant, elles présentent des profils d'expression différents, suggérant des implications dans des processus de développement différents au cours du développement de l'organe : bien que les trois copies de *TaGW2* semblent avoir une implication à la fin du remplissage en réserves du grain de blé, seules les copies A et D sembleraient intervenir au cours de la phase de division du grain. De plus, *TaGW2-D* pourrait également agir durant la phase de maturation du grain. Les travaux de validation fonctionnelle de *TaGW2* par ARN interférence ont montré que la sous-expression des trois copies à la fois impactait de manière significative le rendement chez le blé. De manière intéressante, nous avons observé des réductions significatives des masses fraîches et sèches du grain, de la quantité d'eau dans le grain à maturité et des dimensions du grain, dans les plantes transgéniques, révélant que *TaGW2* est un régulateur positif de la taille et du poids de grain chez le blé tendre, contrairement à *GW2* qui est un régulateur négatif de ces caractères chez le riz (Song *et al.*, 2007). A l'issue de ce travail, nous proposons que les copies *TaGW2* constitueraient un pool d'acteurs positifs du développement du grain de blé tendre en conditions optimales, avec une action présumée au niveau de la dynamique hydrique du grain et/ou de la division cellulaire, au cours de la phase de division du grain ; ainsi qu'au niveau de l'accumulation en réserves amylacées, à la phase de remplissage du grain. Il est encore impossible pour le moment de savoir si toutes les copies de *TaGW2* ou seulement une ou deux en particulier sont responsables de ce contrôle sur le développement du grain. Toutefois, des analyses d'association uni- et multilocales, dans un panel mondial de variétés de blé, ont permis d'identifier du polymorphisme de séquence dans *TaGW2-A* associé

notamment au poids de 1000-grains. De plus, la cartographie génétique de *TaGW2-A* a permis de localiser la copie A de *TaGW2* dans une région comportant plusieurs QTLs pour le rendement et ses composantes.

A la vue de l'ensemble de ces résultats, et des études précédemment parues sur *TaGW2-A* chez le blé tendre, la caractérisation moléculaire et biochimique de la protéine TaGW2-A a été entreprise. Ces travaux de thèse ont permis de montrer que 1) la protéine TaGW2-A est une E3 ubiquitine ligase fonctionnelle *in vitro*, 2) qu'elle présente une localisation nucléoplasmique, nucléolaire et cytoplasmique, régulée par des motifs d'export nucléaire et de rétention nucléolaire situés dans la région N-terminale de la protéine, et 3) que l'activité E3 ligase, par l'intermédiaire du domaine RING, semble importante pour la localisation à la fois nucléolaire et cytoplasmique de la protéine, suggérant une activité de la protéine TaGW2-A dans ces deux compartiments.

Afin de comprendre par quel mécanisme *TaGW2* impacte la taille et le poids du grain de blé, l'enjeu majeur de ce travail de thèse était ensuite d'essayer d'identifier la ou les voie(s) de signalisation dans la(es)quelle(s) la protéine TaGW2 intervient au cours du programme de développement du grain. Le criblage d'une banque ADNc de grains entiers par la technique de double-hybride a permis d'identifier une première liste d'interacteurs protéiques potentiels de TaGW2. De manière intéressante, deux facteurs de transcription appartenant à des familles (NAC et R1R2R3-MYB) décrites chez d'autres espèces dans le contrôle du cycle cellulaire, ainsi que deux protéines impliquées dans la division cellulaire, ont été identifiées dans ce crible. De plus, une protéine RUB1 a été identifiée, suggérant une potentielle double fonction E3 NEDD8 et ubiquitine ligase de TaGW2. Par ailleurs, l'identification d'une enzyme de la voie de synthèse d'un osmoprotectant cellulaire est en accord avec une possible implication de TaGW2-D dans le processus de maturation du grain de blé tendre.

Divergences entre les résultats de ce travail de thèse et les études de Su et al (2011) et Yang et al (2012).

Ces deux dernières années, deux études ont montré la liaison génétique entre *TaGW2-A* et le poids de 1000-grains (Su et al, 2011 ; Yang et al, 2012). Cependant, bien que ces deux études suggèrent un rôle négatif de *TaGW2-A* sur cette composante, tout comme l'homologue *GW2* chez le riz, les deux études divergent quant à la nature du mécanisme causal.

En comparant les trois séquences homéologues de *TaGW2* entre des variétés de blé chinois à petits et gros grains, Su et ses collaborateurs (2011) n'ont trouvé aucune variation dans les séquences codantes, contrairement au mécanisme causal de la perte de fonction *GW2* chez le riz. Par contre, les auteurs ont identifié deux polymorphismes de substitution de bases dans le promoteur de *TaGW2-A* associés à la largeur et au poids de 1000-grains, dans une core-collection de blés chinois. De plus, les auteurs ont établi une corrélation négative entre le niveau d'expression de *TaGW2-A* et la largeur de grains de blé issus de cette core collection, âgés de 10 JAA, suggérant un effet du niveau d'expression de *TaGW2-A* sur la largeur du grain. De leur côté, Yang et ses collaborateurs (2012) n'ont reporté aucune variation dans les régions codantes de *TaGW2-B* et *-D* mais identifié une insertion d'une paire de base dans le huitième exon de *TaGW2-A* d'une variété de blé chinois à gros grain, Lankaodali. Cette insertion entraîne la production d'une protéine non-fonctionnelle, ce qui est similaire au mécanisme de perte de fonction *GW2* chez le riz. Ce marqueur a été retrouvé associé à la largeur du grain et au poids de 1000-grains, dans une population F2 dérivée d'un croisement entre Lankaodali et Chinese Spring. Nous avons montré dans notre étude que la sous-expression des copies *TaGW2* entraîne la production de grains de petite taille. Nos résultats sont cohérents avec l'étude de Su *et al* (2011), puisque leur étude et la nôtre impliquent une modulation de l'expression de *TaGW2* comme mécanisme causal. Cependant, nos résultats sont divergents des études de Su *et al* (2011) et Yang *et al* (2012), puisqu'ils suggèrent que *TaGW2* serait un régulateur positif de ces caractères du grain. Trois hypothèses peuvent être formulées afin d'expliquer cette divergence. La première explication possible est que les deux études précédentes sont focalisées sur la copie A de *TaGW2*, et qu'il n'y a, à l'heure actuelle, aucune information concernant les deux autres copies. L'ARN interférence permet d'éteindre les trois copies de *TaGW2* à la fois. Ainsi, un éventuel effet cumulatif de l'extinction des trois homéologues *TaGW2* tout au long du cycle de développement de la plante peut être observé dans ces conditions. Une seconde explication peut reposer sur la nature des études. Les précédentes études reposent uniquement sur la recherche d'une association génétique entre des marqueurs polymorphes dans *TaGW2-A* et des composantes du rendement. Cependant, les deux études indiquent que *TaGW2-A* est situé dans une région génomique contenant un QTL majeur de rendement, proche du centromère du chromosome 6A (Snape *et al*, 2007 ; revue de Barrero *et al*, 2011). Il ne peut pas être exclu que l'association génétique qu'ils observent soit expliquée par un déséquilibre de liaison entre leurs marqueurs et un gène non identifié encore, ayant un large effet sur les composantes du rendement. De manière intéressante, Li et ses collaborateurs (2010a) ont conduit des études similaires chez le maïs et cherché à identifier

Tableau 17 : Valeurs des largeurs de grain et des taux d'expression relative de *TaGW2* de 14 variétés chinoises de blé, fournies en Figure 5 de l'étude de Su *et al* (2011).

Allèle au SNP (-593)	Variété	Largeur du grain (mm)	Taux d'expression relative de <i>TaGW2</i>
G	Lumai1	3,56	1
G	Mingxian169	3,02	1,25
G	Sanyuehuang	2,96	1,3
G	Mahuaban	2,65	2
G	Baidongmai	2,84	2,5
G	Baimangmai	2,92	3,25
G	Honginmai	2,72	4,25
A	Yannong15	3,35	1,1
A	Zhongyou9507	3,5	1,1
A	Baimaizi	3,2	1,2
A	Beijing8	3,3	1,25
A	Shijiazhuang54	3,18	1,6
A	Lyhan328	3,3	1,9
A	Jiangxizao	2,96	1,9

les allèles favorables de *ZmGW2-CHR4*, c'est-à-dire la contribution relative de chaque allèle parental pour le caractère, en utilisant une analyse d'association et la cartographie de QTL. Leurs deux approches ont généré des résultats contradictoires et les auteurs suggèrent par ailleurs que leur marqueur polymorphe pourrait être en déséquilibre de liaison avec le véritable polymorphisme causal. Une troisième explication possible pourrait être que la cassette RNAi relativement longue que nous avons utilisée pour éteindre l'expression des trois copies de *TaGW2* ait éteint d'autres gènes que *TaGW2*. Bien qu'une recherche dans les bases de données génomiques de blé de l'URGI et CerealsDB n'ait pas révélé de bandes d'identité de séquence de 20nt au minimum avec des séquences codantes connues, nous ne pouvons pas exclure l'hypothèse de l'extinction additionnelle d'autres gènes également impliqués dans le contrôle positif de composantes du rendement.

Su et ses collaborateurs (2011) ont indiqué que le taux d'expression de *TaGW2*, toutes copies confondues, était négativement corrélé à la largeur du grain, dans des grains âgés de 10 JAA, dans 14 variétés de blé chinois (7 variétés avec l'allèle A au marqueur TaGW2_A_R(-593) et 7 variétés avec l'allèle G). Une analyse des données fournies par les auteurs dans la figure 5 de leur étude (les données sont retranscrites dans le Tableau 17) indiquent des coefficients de corrélation de 0,36 et 0,43 entre le taux d'expression relative de *TaGW2* et la largeur du grain dans les variétés avec l'allèle G et les variétés avec l'allèle A, respectivement. De manière opposée, nous avons montré que la réduction de l'expression de *TaGW2*, toutes copies confondues, conduisait à une réduction de la largeur du grain, dans notre étude. Deux explications pourraient être à l'origine de cette divergence. La première possibilité serait que le panel de variétés utilisées par Su et ses collaborateurs (2011) pour leur analyse transcriptionnelle est trop petit pour pouvoir tirer cette conclusion. La deuxième explication est que leur analyse ne porte que sur un seul stade de développement, et qu'aucune information n'est donnée concernant d'une part le profil d'expression de *TaGW2* au cours de l'ensemble du développement du grain et d'autre part le comportement de chacune des copies. De plus, il est relativement difficile de comparer l'effet d'un gène à partir des résultats d'une expérience où le gène a été constitutivement sous-exprimé tout au long du développement de la plante (soit notre analyse) et des résultats du niveau d'expression de ce même gène à un stade de développement donné (soit l'analyse de Su *et al* (2011)). En effet, par exemple, des effets phénotypiques peuvent apparaître à des stades avancés du

développement de la plante ou de l'organe en conséquence de la sous-expression du gène à des stades antérieurs.

Toutefois, quelques résultats des travaux de Su et ses collaborateurs (2011) restent quelque peu troublants. Tout d'abord, comme nous l'avons mentionné dans la discussion du chapitre I, les auteurs indiquent que l'allèle favorable au marqueur TaGW2_A_R(-593) pour le poids de 1000-grains est l'allèle A, tandis que notre étude révèle que ce serait plutôt l'allèle G. De plus, nous avons montré que l'allèle G avait majoritairement été sélectionné dans les variétés de blé à travers le monde. Toutefois, concernant ce résultat, des hypothèses (que nous avons mentionnées dans la discussion du chapitre I) peuvent encore être avancées pour expliquer cette divergence. Cependant, Su et ses collaborateurs ont également indiqué que « le niveau d'expression moyen de *TaGW2* dans les variétés possédant l'allèle G au marqueur était plus élevé que dans les variétés possédant l'allèle A », supportant le fait que l'allèle A soit associé à des largeurs et des poids plus élevés du grain. Or, nous avons analysé les différences dans les données d'expression fournies par Su et ses collaborateurs (Tableau 17) et testé leur significativité par un test de Scheffé, en prenant comme critère de significativité $p < 0,05$. Contrairement à ce que mentionnent Su et ses collaborateurs dans leur analyse, le niveau moyen d'expression de *TaGW2* ne semble pas être significativement différent entre le pool de variétés avec l'allèle G et le pool de variétés avec l'allèle A, puisque nous avons trouvé une *p*-value de 0,12.

Collectivement, l'ensemble de ces résultats semble donc indiquer qu'un caractère quantitatif peut être trop complexe pour être simplement expliqué par une seule approche, et que des conditions expérimentales peuvent vraisemblablement dramatiquement impacter le phénotype obtenu.

Les homologues TaGW2 chez le blé et GW2 chez le riz ont des fonctions divergentes dans le développement du grain.

De manière intéressante, les caractères du grain (poids, largeur et épaisseur) affectés par l'ARN interférence de *TaGW2* dans notre étude sont les mêmes que ceux altérés dans les lignées riz antisense *GW2* produites par Song et ses collaborateurs, et la variété naturelle de riz WY3, qui comporte un variant perte-de-fonction *GW2* (Song *et al*, 2007). Cependant, *TaGW2* agit de manière positive sur ces caractères, tandis que *GW2* agit négativement. De plus, dans les plantes transgéniques RNAi *TaGW2*, nous n'avons pas observé de différences

significatives dans le nombre de grains par épi et dans le nombre d'épis par plante par rapport aux plantes contrôles, tandis que la perte de fonction *GW2* chez le riz réduit de 29,9% le nombre de grains par épi et d'environ 27% le nombre de panicule par plante. Il peut être important de souligner également que chez le riz, *GW2* régule le nombre de cellules dans les enveloppes externes du grain mais n'a aucun effet sur le nombre de cellules dans l'albumen. De manière contrastée, nous avons vu que *TaGW2* peut réguler le nombre de cellules de l'albumen, tout au moins dans la lignée transgénique T105. Les deux homologues *GW2* et *TaGW2* semblent donc fonctionner différemment dans le développement du grain de riz et de blé, respectivement. Cette divergence de fonction entre deux gènes homologues est très inattendue étant donné que des gènes homologues possèdent souvent des fonctions conservées entre des espèces proches (pour exemple Fan *et al*, 2006 ; Li *et al*, 2010b).

Bien que le phénomène soit encore peu décrit dans la littérature, toutefois, quelques études reportent des divergences de fonctions entre « orthologues ». Des relations gène-phénotype peuvent être altérées de manière très extrême entre des espèces, montrant une implication d'un gène dans des processus de développement complètement différents suivant l'espèce considérée, comme par exemple le gène *GSK3* chez la levure, les insectes, la grenouille et la souris (Kassir *et al*, 2006 ; Liu *et al*, 2007a). Gharib et Robinson-Rechavi (2011) ont recensé dans leur revue plusieurs exemples de divergences de fonction de gènes entre l'Homme et la souris. Les auteurs reportent notamment divers mécanismes pouvant être à l'origine de cette divergence de fonction : des différences dans la régulation de l'expression des gènes « orthologues » (divergence de profil d'expression temporelle, spatiale), différences dans le nombre de copies du gène considéré, mais également des différences dans la localisation subcellulaire des protéines « orthologues ». Fait très intéressant, il apparaîtrait que les gènes de ménage orthologues, exprimés de manière constitutive, seraient plus sujets à divergence de profil d'expression que les gènes tissu-spécifiques, entre l'Homme et la souris. De plus, Liao et Zhang (2011) ont montré que plus de 20% des gènes essentiels chez l'Homme se retrouvent être des gènes non-essentiels chez le souris. De plus, même une conservation d'essentialité d'un gène n'assure pas une conservation de phénotype entre les gènes, comme par exemple dans le cas des gènes *Adamts2*, *Acox1* et *Fancg* qui sont des gènes essentiels chez l'Homme et la souris mais qui présentent des phénotypes différents lorsqu'ils sont mutés chez ces deux espèces (Fan *et al*, 1996 ; Jimenez-Snachez *et al*, 2001 ; Yang *et al*, 2001b ; Suzuki *et al*, 2002 ; Liao et Zhang, 2011).

De manière intéressante, Distelfeld et ses collaborateurs (2012) ont reporté très récemment une divergence de fonction entre des homologues du gène *GPC*, *GPC-B2* et

Os07g37920, chez le blé et le riz respectivement. Bien que ces gènes soient très proches phylogénétiquement et se retrouvent dans des régions colinéaires dans les deux génomes, ils participent à des programmes de développement complètement différents. Les gènes *GPC*, dont *GPC-B2*, contrôlent la sénescence foliaire chez le blé, tandis que le gène *Os07g37920* contrôle la déhiscence des anthères chez le riz. De plus, les plantes transgéniques *TaGPC-B* ne sont pas affectées pour la fertilité et la déhiscence des anthères ; et les plantes transgéniques *Os07g37920* ne sont pas affectées dans la sénescence foliaire. Par ailleurs, il a été remarqué que le profil d'expression spatio-temporel des gènes « orthologues » différait entre les deux espèces. Ohba et ses collaborateurs (2000) ont également montré des divergences de réponse à des stimuli environnementaux entre des homologues du gène *WPK4*, un gène codant une protéine kinase impliquée dans la signalisation par les sucres, chez le blé, le maïs et le riz.

Il peut être imaginé que les programmes de développement du grain, et plus particulièrement la détermination du potentiel de taille du grain, ne soient pas strictement les mêmes entre le riz et le blé tendre, constituant ainsi une des pistes d'explication à la divergence de fonction de *GW2* et *TaGW2*. En effet, par exemple, les contraintes physiques à l'établissement de la taille du grain ne sont pas strictement établies par les mêmes organes entre les deux espèces. Chez le riz, les glumelles sont non-adhérentes au grain mais soudées entre elles, et constituent ainsi une enveloppe rigide autour du grain qui agit directement comme une barrière mécanique à la croissance de celui-ci. Chez le blé, les glumelles sont non-soudées entre elles et ne renferment pas fermement le grain. Elles ne représentent donc pas une barrière physique à la croissance du grain, comme chez le riz. Par conséquent, il pourrait être admis que des glumelles plus larges chez le blé ne conduisent pas à la production de grains plus larges, comme ce qui se passe chez le riz, mais qu'une augmentation intrinsèque du volume du grain soit un moyen plus probable d'y arriver. Il est intéressant de rappeler que *GW2*, chez le riz, n'a aucune action sur le nombre de cellules du grain à proprement parler, mais une action sur le nombre de cellules dans les glumelles. Dans notre étude, nous n'avons malheureusement pas étudié l'effet de l'ARN interférence de *TaGW2* dans les glumelles du grain, mais nous avons observé un effet de *TaGW2* sur le volume du grain et sur le nombre de cellules de l'albumen. La divergence de fonction entre *GW2* et *TaGW2* pourrait déjà résider dans cette différence du lieu d'action.

De plus, comme Gharib et Robinson-Rechavi (2011) l'ont mentionné, une divergence dans la localisation subcellulaire de protéines « orthologues » peut être à l'origine de la

divergence de fonction. Rappelons que nous avons montré au cours de notre étude que la protéine TaGW2-A présentait une localisation nucléoplasmique, nucléolaire et cytoplasmique ; tandis que Song *et al* (2007) ont montré que la protéine GW2 était cytoplasmique. Bien que les divers signaux que nous avons identifiés dans la protéine TaGW2-A au cours de cette étude soient conservés dans la séquence primaire de GW2, nous n'avons toujours pas identifié de signaux responsables de l'import nucléaire de la protéine. Une différence au niveau d'un de ces signaux pourrait expliquer les différences de localisation subcellulaire des deux protéines homologues. *In fine*, une divergence de localisation subcellulaire des protéines GW2 et TaGW2 pourraient être à l'origine de la divergence de fonction.

Bien que les deux protéines partagent 87% d'identité de séquence totale en acides aminés, il peut également être envisagé que les deux protéines « orthologues » aient des fonctions divergentes du fait d'interaction avec des protéines différentes. Les deux séquences primaires sont hautement conservées en N-terminale (93,8% d'identité de séquence entre les acides aminés 1 à 180) supportant par ailleurs une conservation de fonction E3 ligase, mais divergent plus en C-terminale (82% d'identité des acides aminés 181 à 424). Un ou des domaines d'interactions avec des protéines cibles pourraient se situer dans cette région, comme l'avaient par ailleurs déjà suggéré Song et ses collaborateurs (2007), et des substitutions en acides aminés pourraient alors venir altérer des résidus clés pour ces interactions avec des régulateurs de voies de signalisation. Il ne peut pas non plus être exclu que les protéines GW2 et TaGW2 ciblent des protéines clé similaires, mais que ce soient ces protéines qui aient des fonctions divergentes entre le blé et le riz.

Perspectives à ce travail de thèse

Un des enjeux majeurs de ce travail est la détermination de la ou les voies de signalisation dans la(es)quelle(s) intervient la protéine TaGW2. La perspective directe de ces travaux consisterait donc en la confirmation des interactions protéine-protéine des partenaires protéiques potentiels de TaGW2 identifiés au cours du crible de la banque ADNc par double-hybride. Après une première élimination des faux-positifs par le test double-hybride 2 à 2, les interactions potentielles seront testées par co-immunoprécipitation et par BiFC afin d'être validées. Il serait ensuite intéressant de regarder si le ou les interacteurs protéiques de TaGW2 sont des cibles de TaGW2, c'est-à-dire sont modifiés par TaGW2, ou sont simplement des partenaires nécessaires à l'action de TaGW2. Dans le cas d'une modification de cet interacteur par TaGW2 par ubiquitination, il serait intéressant de regarder si la modification est un étiquetage par des chaînes de polyubiquitine en K48, conduisant à la dégradation de la protéine ciblée, ou consiste en une monoubiquitination ou l'apposition de chaînes non-conventionnelles de polyubiquitine, nécessaires par exemple à la stabilisation du substrat ou l'augmentation de son activité, ou encore à sa localisation subcellulaire. De plus, dans le cas où l'interaction de TaGW2 et la protéine NEDD8 serait confirmée, il pourrait être intéressant de tester si TaGW2 est effectivement capable de neddylation. Ceci pourrait être initié par un test d'auto-neddylation *in vitro*, voire poursuivi par un test de neddylation *in vitro* des interacteurs protéiques confirmés de TaGW2.

Afin d'essayer de déterminer l'implication de chaque copie de *TaGW2* dans l'établissement du programme de développement du grain, il serait intéressant de produire des plantes transgéniques sur-exprimant indépendamment chacune des copies. Des grains de blé transgéniques T1 sur-exprimant la copie A de *TaGW2* sont d'ores et déjà disponibles, et un phénotypage de ces plantes transgéniques permettrait peut-être d'avoir une idée de la fonction de *TaGW2-A* au cours du développement. Ceci permettrait peut-être également de clarifier les divergences de résultats observés entre notre étude et les études de Su *et al* (2011) et Yang *et al* (2012). Cependant, il est nécessaire de rappeler qu'un phénotype peut ne pas être visible en conséquence de la surexpression de *TaGW2-A* du fait d'éventuelles compensations de fonction souvent montrées entre des homéologues, chez des espèces polyploïdes.

La séquence promotrice de *TaGW2-D* est disponible dans la base de données de l'URGI. Malheureusement, il n'y a pour l'instant aucune information disponible concernant

les séquences promotrices des deux autres copies de *TaGW2*. De manière intéressante, l'analyse de 2004 pb de la séquence du promoteur de *TaGW2-D*, en amont du site d'initiation de la transcription, révèle de nombreuses boîtes de régulation, dont un nombre conséquent de boîtes de signalisation à l'acide abscissique (ABA ; Annexe 6). La phytohormone ABA est un régulateur clé de très nombreux processus de développement en conditions optimales de croissance et de réponses adaptatives des plantes à des stress environnementaux (Cutler *et al*, 2010 ; Fujita *et al*, 2011 ; Wilkinson *et al*, 2012). En conditions optimales, dans le développement du grain chez les céréales en particulier, les niveaux d'ABA ont été décrits intimement liés au remplissage en amidon du grain, et à la préparation à la dessiccation du grain au cours de la phase de maturation, mais également, plus récemment chez l'orge aux processus de cellularisation, division cellulaire et endoreduplication au cours de la phase précoce du développement du grain (Sreenivasulu *et al*, 2010 ; Seiler *et al*, 2011 ; Zhu *et al*, 2011). De plus, en conditions de stress osmotiques, tels qu'un stress hydrique ou salin, l'augmentation des niveaux d'ABA induit l'expression de gènes de réponse et de tolérance à la dessiccation. Il pourrait être intéressant de rechercher dans un premier temps, si l'action de *TaGW2* sur la détermination du potentiel de poids du grain, en conditions optimales, passe par une signalisation à l'ABA. De manière intéressante, plusieurs éléments abordés au cours de ce travail de thèse pourraient également laisser penser qu'il existe un lien entre *TaGW2* et les stress abiotiques. En effet, l'expression des différentes copies de *TaGW2* semble être modulée par un « stress thermique » au cours de la phase de remplissage et des boîtes de signalisation à l'ABA et au stress hydrique ont été identifiées dans le promoteur de *TaGW2-D*. De plus, plusieurs interacteurs potentiels de *TaGW2* que nous avons identifiés à l'issue du crible double-hybride ont été précédemment décrits dans la régulation de stress abiotiques dont le stress hydrique, tels que les FT NAC et MYB, l'osmoprotectant bétaine-glycine, ou encore une OLP (osmotin-like protein). Une modulation de l'action de *TaGW2* sur le potentiel de poids du grain par des facteurs abiotiques pourrait par ailleurs expliquer que *TaGW2-A* soit retrouvé dans une région génomique contenant des QTL de composantes du rendement instables entre les environnements.

REFERENCES BIBLIOGRAPHIQUES

-A-

- Acevedo E, Silva P, Silva H. 2002. Growth and wheat physiology, development. In: Bread wheat: improvement and production. Edited by Curtis BC, Rajaram S, Gomez Macpherson H: Food and Agriculture Organization of the United Nations (FAO); 2002: 567 pg
- Albert TK, Hanzawa H, Legtenberg YI, de Ruwe MJ, van den Heuvel FA, Collart MA, Boelens R, Timmers HT. 2002. Identification of a ubiquitin-protein ligase subunit within the CCR4-NOT transcription repressor complex. *EMBO Journal* 21:355-364.
- Altenbach SB, Dupont FM, Kothari KM, Chan R, Johnson EL, Lieu D. 2003. Temperature, water and fertilizer influence the timing of key events during grain development in a US spring wheat. *Journal of Cereal Science* 37:9-20.
- Altenbach SB. 2012. New insights into the effects of high temperature, drought and post-anthesis fertilizer on wheat grain development. *Journal of Cereal Science* 56:39-50.
- An DG, Su JY, Liu QY, Zhu YG, Tong YP, Li JM, Jing RL, Li B, Li ZS. 2006. Mapping QTL for nitrogen uptake in relation to the early growth of wheat (*Triticum aestivum* L.). *Plant Soil* 284:73-84.
- Andersen JS, Lam YW, Leung AK, Ong SE, Lyon CE, Lamond AI, Mann M. 2005. Nucleolar proteome dynamics. *Nature* 433:77-83.
- Andersson A, Johansson E, Oscarson P. 2004. Post-anthesis nitrogen accumulation and distribution among grains in spring wheat spikes. *Journal of Agricultural Science* 142:525-533.
- Ang LH, Chattopadhyay S, Wei N, Oyama T, Okada T, Okada K, Batschauer A, Deng XW. 1998. Molecular interaction between COP1 and HY5 defines a regulatory switch for light control of Arabidopsis development. *Molecular Cell* 1:213-222.
- Antoine C, Lullien-Pellerin V, Abecassis J, Rouau X. 2002. Nutritional interest of the wheat seed aleurone layer. *Sciences des Aliments* 22:545-556
- Arabi A, Rustum C, Hallberg E, Wright AP. 2003. Accumulation of c-Myc and proteasomes at the nucleoli of cells containing elevated c-Myc protein levels. *Journal of Cell Science* 116:1707-1717.
- Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles ER, Qian Q, Kitano H, Matsuoka M. 2005. Cytokinin oxidase regulates rice grain production. *Science* 309:741-745.

-B-

- Balfourier F, Roussel V, Strelchenko P, Exbrayat-Vinson F, Sourdille P, Boutet G, Koenig J, Ravel C, Mitrofanova O, Beckert M, Charmet G. 2007. A worldwide bread wheat core collection restricted to a full 384 deep well storage plate. *Theoretical and Applied Genetics* 114:1265-1275.
- Barrero RA, Bellgard M, Zhang X. 2011. Diverse approaches to achieving grain yield in wheat. *Functional & Integrative Genomics* 11:37-48.
- Barron C, Surget A, Rouau X. 2007. Relative amounts of tissues in mature wheat (*Triticum aestivum* L.) grain and their carbohydrate and phenolic acid composition. *Journal of Cereal Science* 45:88-96
- Baulcombe D. 2004. RNA silencing in plants. *Nature* 431:356-363.

- Bechtel DB, Zayas I, Kaleikau L, Pomeranz Y. 1990. Size-distribution of wheat-starch granules during endosperm development. *Cereal Chemistry* 67:59-63
- Berger F. 1999. Endosperm development. *Current Opinion in Plant Biology* 2:28-32.
- Birbach A, Bailey ST, Ghosh S, Schmid JA. 2004. Cytosolic, nuclear and nucleolar localization signals determine subcellular distribution and activity of the NF-kappaB inducing kinase NIK. *Journal of Cell Science* 117:3615-3624.
- Blilou I, Frugier F, Folmer S, Serralbo O, Willemsen V, Wolkenfelt H, Eloy NB, Ferreira PC, Weisbeek P, Scheres B. 2002. The Arabidopsis HOBbit gene encodes a CDC27 homolog that links the plant cell cycle to progression of cell differentiation. *Genes & Development* 16:2566-2575.
- Bogorad L, Gubbins EJ, Krebbers E, Larrinua IM, Mulligan BJ, Muskavitch KMT, Orr EA, Rodermel SR, Schantz R, Steinmetz AA, De Vos G, Ye YK. 1983. Cloning and physical mapping of maize plastid genes. *Methods in Enzymology* 97:524-554.
- Boisvert FM, van Koningsbruggen S, Navascués J, Lamond AI. 2007. The multifunctional nucleolus. *Nature Reviews in Molecular Cell Biology* 8:574-585.
- Börner A, Schumann E, Fürste A, Cöster H, Leithold B, Röder S, Weber E. 2002. Mapping of quantitative trait loci determining agronomic important characters in hexaploid wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* 105:921-936.
- Bordes J, Branlard G, Oury FX, Charmet G, Balfourier F. 2008. Agronomic characteristics, grain quality and flour rheology of 372 bread wheats in a worldwide core collection. *Journal of Cereal Science* 48:569-579.
- Boulton SJ. 2006. Cellular functions of the BRCA tumour-suppressor proteins. *Biochemical Society Transactions* 34:633-645.
- Boyeldieu J. 1997. Blé tendre, *Techniques Agricoles-Fascicule n°2020*, Paris.
- Bradbury D, MacMasters MM, Cull IM. 1956. Structure of the mature wheat kernel. Microscopic structure of pericarp, seed coat, and other coverings of the endosperm and germ of hard red winterwheat. *Cereal Chemistry* 33:342-360
- Breseghele F, Sorrells ME. 2007. QTL analysis of kernel size and shape in two hexaploid wheat mapping population. *Field Crops Research* 101:172-179
- Briarty LG, Hughes CE, Evers AD. 1979. Developing endosperm of wheat – a stereological analysis. *Annals of Botany* 44:641-658.
- Brocklehurst PA. 1977. Factors controlling grain weight in wheat. *Nature* 266:348-349.
- Brooks A, Jenner CF, Aspinall D. 1982. Effects of water deficit on endosperm starch granules and on grain physiology of wheat and barley. *Australian Journal of Plant Physiology* 9:423-436.
- Brooks SA. 2010. Functional interactions between mRNA turnover and surveillance and the ubiquitin proteasome system. *RNA* 1; doi: 10.1002/wrna.11.
- Brückner A, Polge C, Lentze N, Auerbach D, Schlattner U. 2009. Yeast Two-Hybrid, a Powerful Tool for Systems Biology. *International Journal of Molecular Sciences* 10:2763-2788.

Brzovic PS, Keefe JR, Nishikawa H, Miyamoto K, Fox D, Fukuda M, Ohta T, Klevit R. 2003. Binding and recognition in the assembly of an active BRCA1/BARD1 ubiquitin-ligase complex. *Proceedings of National Academy of Sciences of United States of America* 100:5646-5651.

Bu Q, Li H, Zhao Q, Jiang H, Zhai Q, Zhang J, Wu X, Sun J, Xie Q, Wang D, Li C. 2009. The Arabidopsis RING finger E3 ligase RHA2a is a novel positive regulator of abscisic acid signaling during seed germination and early seedling development. *Plant Physiology* 150:463-481.

-C-

Calderini DF, Abeledo LG, Savin R, Slafer GA. 1999a. Effect of temperature and carpel size during pre-anthesis on potential grain in wheat. *Journal of Agricultural Science, Cambridge* 132:453-459.

Calderini DF, Abeledo LG, Savin R, Slafer GA. 1999b. Final grain weight in wheat as affected by short periods of high temperature during pre- and post-anthesis under field conditions. *Australian Journal of Plant Physiology* 26:453-458.

Calderini DF, Reynolds MP. 2000. Changes in grain weight as a consequence of de-graining treatments at pre- and post-anthesis in synthetic hexaploid lines of wheat. *Australian Journal of Plant Physiology* 27:183-191.

Calderini DF, Ortiz-Monasterio I. 2003. Grain position affects grain macronutrient and micronutrient concentrations in wheat. *Crop science* 43:141-151

Caley CY, Duffus CM, Jeffcoat B. 1990. Effects of elevated temperature and reduced water uptake on enzymes of starch synthesis in developing wheat grains. *Australian Journal of Plant physiology* 17:431-439.

Callis J, Carpenter T, Sun CW, Vierstra RD. 1995. Structure and evolution of genes encoding polyubiquitin and ubiquitin-like proteins in Arabidopsis thaliana ecotype Columbia. *Genetics* 139:921-939.

Campbell KJ, Bergman CJ, Gualberto DG, Anderson JA, Giroux MJ, Hareland G, Fulcher RG, Sorrells ME, Finney L. 1999. Quantitative Trait Loci Associated with Kernel Traits in a Soft \times Hard Wheat Cross. *Crop Science* 39:1184-1195

Cao Y, Dai Y, Cui S, Ma L. 2008. Histone H2B Monoubiquitination in the Chromatin of FLOWERING LOCUS C Regulates Flowering Time in Arabidopsis. *The Plant Cell* 20: 2586-2602.

Capron A, Okresz L, Genschik P. 2003. First glance at the plant APC/C, a highly conserved ubiquitin-protein ligase. *Trends in Plant Science* 8:83-89.

Capron D, Mouzeyar S, Boulaflous A, Girousse C, Rustenholz C, Laugier C, Paux E, Bouzidi MF. 2012. Transcriptional profile analysis of E3 ligase and hormone-related genes expressed during wheat grain development. *BMC Plant Biology* 14:12-35.

Catalá R, Medina J, Salinas J. 2011. Integration of low temperature and light signaling during cold acclimation response in Arabidopsis. *Proceedings of the National Academy of Sciences of United States of America* 108:16475-16480.

Chae E, Tan QK, Hill TA, Irish VF. 2008. An Arabidopsis F-box protein acts as a transcriptional co-factor to regulate floral development. *Development* 135:1235-1245.

Charmet G, Robert N, Branlard G, Linossier L, Martre P, Tribot E. 2005. Genetic analysis of dry matter and nitrogen accumulation and protein composition in wheat kernels. *Theoretical and Applied Genetics* 111:540-550.

Charmet G, Masood-Quraishi U, Ravel C, Romeuf I, Balfourier F, Perretant MR, Joseph JL, Rakszegi M, Guillon F, Sado PE, Bedő Z, Saulnier L. 2009. Genetics of dietary fibre in bread wheat. *Euphytica* 170:155-168.

Chen M, Rockel T, Steinweger G, Hemmerich P, Risch J, von Mikecz A. 2002. Subcellular recruitment of fibrillarin to nucleoplasmic proteasomes: implications for processing of a nucleolar autoantigen. *Molecular Biology of the Cell* 13:3576-3587.

Chen H, Shen Y, Tang X, Yu L, Wang J, Guo L, Zhang Y, Zhang H, Feng S, Strickland E, Zheng N, Deng XW. 2006. Arabidopsis CULLIN4 forms an E3 Ubiquitin Ligase with RBX1 and the CDD complex in mediating light control of development. *The Plant Cell* 18:1991-2004.

Chen TH, Murata N. 2011. Glycinebetaine protects plants against abiotic stress: mechanisms and biotechnological applications. *Plant, Cell & Environment* 34:1-20.

Chojacki AJS, Bayliss MW, Gale MD. 1986. Cell production and DNA accumulation in the wheat endosperm, and their association with grain weight. *Annals of Botany* 58:809-817.

Christians MJ, Gingerich DJ, Hansen M, Binder BM, Kieber JJ, Vierstra RD. 2009. The BTB ubiquitin ligases ETO1, EOL1 and EOL2 act collectively to regulate ethylene biosynthesis in Arabidopsis by controlling type-2 ACC synthase levels. *Plant Journal* 57:332-345.

Christensen DE, Brzovic PS, Klevit RE. 2007. E2-BRCA1 RING interactions dictate synthesis of mono- or specific polyubiquitin chain linkages. *Nature Structural and Molecular Biology* 14:941-948.

Clough RC, Jordan-Beebe ET, Lohman KN, Marita JM, Walker JM, Gatz C, Vierstra RD. 1999. Sequences within both the N- and C-terminal domains of phytochrome A are required for PFR ubiquitination and degradation. *Plant Journal* 17:155-167.

Cossegal M, Vernoud V, Depege N, Rogowsky PM. 2007. The embryo surrounding region. In: OA Olsen, ed, *Endosperm*, Vol 8. Springer-Verlag, Berlin/Heidelberg, pp 57-71.

Crocco CD, Holm M, Yanovsky MJ, Botto JF. 2010. AtBBX21 and COP1 genetically interact in the regulation of shade avoidance. *Plant Journal* 64:551-562.

Cutler SR, Rodriguez PL, Finkelstein RR, Abrams SR. 2010. Absciscic acid: emergence of a core signaling network. *Annual Review of Plant Biology* 61:651-679.

-D-

Dai X, Xu Y, Ma Q, Xu W, Wang T, Xue Y, Chong K. 2007. Overexpression of an R1R2R3 MYB gene, OsMYB3R-2, increases tolerance to freezing, drought, and salt stress in transgenic Arabidopsis. *Plant Physiology* 143:1739-1751.

De Leonardis AM, Marone D, Mazzucotelli E, Neffar F, Rizza F, Di Fonzo N, Cattivelli L, Mastrangelo AM. 2007. Durum wheat genes up-regulated in the early phase of cold stress are

modulated by drought in a developmental and genotype dependent manner. *Plant Science* 172:1005–1016.

De Vienne D. 1995. Les marqueurs moléculaires en génétique et biotechnologies végétales. INRA Editions

Del Pozo JC, Timpte C, Tan S, Callis J, Estelle M. 1998. The ubiquitin-related protein RUB1 and auxin response in *Arabidopsis*. *Science* 280:1760-1763.

Del Pozo JC, Dharmasiri S, Hellmann H, Walker L, Gray WM, Estelle M. 2002. AXR1-ECR1-dependent conjugation of RUB1 to the *Arabidopsis* Cullin AtCUL1 is required for auxin response. *The Plant Cell* 14:421-433.

Deng XW, Caspar T, Quail PH. 1991. *cop1*: a regulatory locus involved in light-controlled development and gene expression in *Arabidopsis*. *Genes & Development* 5: 1172-1182.

Deng S, Wu X, Wu Y, Zhou R, Wang H, Jia J, Liu S. 2011. Characterization and precise mapping of a QTL increasing spike number with pleiotropic effects in wheat. *Theoretical and Applied Genetics* 122:281-289.

Deshaies RJ, Joazeiro CAP. 2009. RING domain E3 ubiquitin ligases. *Annual reviews of Biochemistry* 78:399-434.

Dholakia BB, Ammiraju JSS, Singh H, Lagu MD, Röder MS, Rao VS, Dhaliwal HS, Ranjekar PK, Gupta VS, Weber WE. 2003. Molecular marker analysis of kernel size and shape in bread wheat. *Plant Breeding* 122:392–395

Distelfeld A, Pearce SP, Avni R, Scherer B, Uauy C, Piston F, Slade A, Zhao R, Dubcovsky J. 2012. Divergent functions of orthologous NAC transcription factors in wheat and rice. *Plant Molecular Biology* 78:515-524.

Downes BP, Stupar RM, Gingerich DJ, Vierstra RD. 2003. The HECT ubiquitin protein ligase (UPL) family in *Arabidopsis*: UPL3 has a specific role in trichome development. *Plant Journal* 35:729–742.

Drea S, Leader DJ, Arnold BC, Shaw P, Dolan L, Doonan JH. 2005. Systematic spatial analysis of gene expression during wheat caryopsis development. *Plant Cell* 17:2172-2185.

Du H, Zhang L, Liu L, Tang XF, Yang WJ, Wu YM, Huang YB, Tang YX. 2009. Biochemical and molecular characterization of plant MYB transcription factor family. *Biochemistry* 74:1–11.

Dubcovsky J, Dvorak J. 2007. Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science* 316:1862-1866.

Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L. 2010. MYB transcription factors in *Arabidopsis*. *Trends in Plant Science* 15:573–581.

Dupont FM, Altenbach SB. 2003. Molecular and biochemical impacts of environmental factors on wheat grain development and protein synthesis. *Journal of Cereal Science* 38:133-146.

Dyachok J, Zhu L, Liao F, He J, Huq E, Blancaflor EB. 2011. SCAR mediates light-induced root elongation in *Arabidopsis* through photoreceptors and proteasomes. *Plant Cell* 23:3610-3326.

-E-

- Earley KW, Haag JR, Pontes O, Opper K, Juehne T, Song K, Pikaard CS. 2006. Gateway-compatible vectors for plant functional genomics and proteomics. *The Plant Journal* 45:616-629.
- Egli DB. 1998. Seed biology and the yield of grain crops. CAB International, New York.
- Emmott E, Hiscox JA. 2009. Nucleolar targeting: the hub of the matter. *EMBO Reports*. 10:231-238.
- Ernoult-Lange M, Wilczynska A, Harper M, Aigueperse C, Dautry F, Kress M, Weil D. 2009. Nucleocytoplasmic traffic of CPEB1 and accumulation in Crml nucleolar bodies. *Molecular Biology of the Cell* 20:176-187.
- Evers AD, Bechtel DB. 1988. Microscopic structure of the wheat grain. *Wheat: chemistry and technology*. Volume I, 47-95
- Evers AD, Blakeney AB, O'Brien L. 1999. Cereal structure and composition. *Australian Journal of Agricultural Research* 50:629-650
- Evers T, Millar S. 2002. Cereal grain structure and development : some implication for quality. *Journal of Cereal Science* 36:261-284

-F-

- Fan CY, Pan J, Chu R, Lee D, Kluckman KD, Usuda N, Singh I, Yeldandi AV, Rao MS, Maeda N, Reddy JK. 1996. Hepatocellular and hepatic peroxisomal alterations in mice with a disrupted peroxisomal fatty acyl-coenzyme A oxidase gene. *Journal of Biological Chemistry* 271:24698–24710.
- Fan CC, Xing YZ, Mao HL, Lu TT, Han B, Xu CG, Li XH, Zhang QF. 2006. GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theoretical and Applied Genetics* 112:1164–1171.
- Farrás R, Ferrando A, Jásik J, Kleinow T, Okrész L, Tiburcio A, Salchert K, del Pozo C, Schell J, Koncz C. 2001. SKP1-SnRK protein kinase interactions mediate proteasomal binding of a plant SCF ubiquitin ligase. *EMBO Journal* 20:2742–2756.
- Favory JJ, Stec A, Gruber H, Rizzini L, Oravecz A, Funk M, Albert A, Cloix C, Jenkins GI, Oakeley EJ, Seidlitz HK, Nagy F, Ulm R. 2009. Interaction of COP1 and UVR8 regulates UV-B-induced photomorphogenesis and stress acclimation in *Arabidopsis*. *EMBO Journal* 28:591-601.
- Feillet P. 2000. Le grain de blé: Composition et utilisation. INRA Editions
- Feldman M. 1995. Wheats. In: SmarttJ, Simmonds NW, eds. *Evolution of crops plants* Harlow, UK: Logman Scientific and Technical, 185-192
- Feller A, Machemer K, Braun EL, Grotewold E. 2011. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *The Plant Journal* 66:94–116.
- Finley D, Bartel B, Varshavsky A. 1989. The tails of ubiquitin precursors are ribosomal proteins whose fusion to ubiquitin facilitates ribosome biogenesis. *Nature* 338: 394–401.

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391:806-811.

Fisher RA, Aguilar I, Laing DR. 1977. Post-anthesis sink size in a high-yielding dwarf wheat: yield response to grain number. *Australian Journal of Agricultural Research* 28:165-175.

Fischer R. 2007. Understanding the physiological basis of yield potential in wheat. *Journal of Agricultural Science* 145:99-113.

Fleury D, Himanen K, Cnops G, Nelissen H, Boccardi TM, Maere S, Beemster GT, Neyt P, Anami S, Robles P, Micol JL, Inzé D, Van Lijsebettens M. 2007. The *Arabidopsis thaliana* homolog of yeast BRE1 has a function in cell cycle regulation during early leaf and root growth. *Plant Cell* 19:417-432.

Fluckiger R, De Caroli M, Piro G, Dalessandro G, Neuhaus JM, Di Sansebastiano GP. 2003. Vacuolar system distribution in *Arabidopsis* tissues, visualized using GFP fusion proteins. *Journal of Experimental Botany* 54:1577-1584.

Fornerod M, Ohno M, Yoshida M, Mattaj JW. 1997. CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* 90:1051-1060.

Fu D, Uauy C, Blechl A, Dubcovsky J. 2007. RNA interference for wheat functional gene analysis. *Transgenic Research* 16:689-701.

Fu D, Uauy C, Distelfeld A, Blechl A, Epstein L, Chen X, Sela H, Fahima T, Dubcovsky J. 2009. A kinase-START gene confers temperature-dependent resistance to wheat stripe rust. *Science* 323:1357-1360.

Fujita Y, Fujita M, Shinozaki K, Yamaguchi-Shinozaki K. 2011. ABA-mediated transcriptional regulation in response to osmotic stress in plants. *Journal of Plant Research*. 124:509-525.

Fukuda M, Gotoh I, Gotoh Y, Nishida E. 1996 Cytoplasmic localization of mitogen-activated protein kinase kinase directed by its NH₂-terminal, leucine-rich short amino acid sequence, which acts as a nuclear export signal. *Journal of Biological Chemistry* 271, 20024-20028.

Fukuda M, Asano S, Nakamura T, Adachi M, Yoshida M, Yanagida M, Nishida E. 1997. CRM1 is responsible for intracellular transport mediated by the nuclear export signal *Nature* 390:308-311.

-G-

Gao T, Wu Y, Zhang Y, Liu L, Ning Y, Wang D, Tong H, Chen S, Chu C, Xie Q. 2011. OsSDIR1 overexpression greatly improves drought tolerance in transgenic rice. *Plant Molecular Biology* 76:145-156.

Garcia D, Saingery V, Chambrier P, Mayer U, Jürgens G, Berger F. 2003. *Arabidopsis* haiku mutants reveal new controls of seed size by endosperm. *Plant Physiology* 131:1661-1670.

Gavin AC, Bosche M, Krause R et al. 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415:141-147

- Gegas VC, Nazari A, Griffiths S, Simmonds J, Fish L, Orford S, Sayers L, Doonan JH, Snape JW. 2010. A Genetic Framework for Grain Size and Shape Variation in Wheat. *The Plant Cell* 22:1046–1056.
- Geil C, Schwab M, Seufert W. 2008. A nucleoluslocalized activator of Cdc14 phosphatase supports rDNA segregation in yeast mitosis. *Current Biology* 18:1001–1005.
- Gervais L, Deddryver F, Morlais JY, Bodusseau V, Negre S, Bilous M, Groos C, Trottet M. 2003. Mapping of quantitative trait loci for field resistance to Fusarium head blight in an European winter wheat. *Theoretical and Applied Genetics* 106:961-970.
- Gharib W, Robinson-Rechavi M. 2011. When orthologs diverge between human and mouse. *Briefings in bioinformatics* 12:436-441.
- Glickman MH, Raveh D. 2005. Proteasome plasticity. *FEBS Letters* 579:3214-3223.
- Gomord V, Faye L. 1996. Signals and mechanisms involved in intracellular transport of secreted proteins in plants. *Plant Physiology and Biochemistry* 34:165-181.
- Greve K, La Cour T, Jensen MK, Poulsen FM, Skriver K. 2003. Interactions between plant RING-H2 and plant-specific NAC (NAM/ATAF1/2/CUC2) proteins: RING-H2 molecular specificity and cellular localization. *The Biochemical Journal* 371:97-108.
- Groos C, Robert N, Bervas E, Charmet G. 2003. Genetic analysis of grain protein-content, grain yield and thousand-kernel weight in bread wheat. *Theoretical and Applied Genetics* 106:1032-1040.
- Guerra D, Mastrangelo AM, Lopez-Torrejon G, Marzin S, Schweizer P, Stanca AM, del Pozo JC, Cattivelli L, Mazzucotelli E. 2012. Identification of a protein network interacting with TdRF1, a wheat RING ubiquitin ligase with a protective role against cellular dehydration. *Plant Physiology* 158:777-789.
- Guo Z, Song E, Ma S, Wang X, Gao S, Shao C, Hu S, Jia L, Tian R, Xu T, Gao Y. 2012. Proteomics Strategy to Identify Substrates of LNX, a PDZ Domain-containing E3 Ubiquitin Ligase. *Journal of Proteome Research* 11:4847-4862
- Gupta R, Kus B, Fladd C, Wasmuth J, Tonikian R, Sidhu S, Krogan NJ, Parkinson J, Rotin D. 2007. Ubiquitination screen using protein microarrays for comprehensive identification of Rsp5 substrates in yeast. *Molecular Systems Biology* 3:116

-H-

- Haasen D, Köhler C, Neuhaus G, Merkle T. 1999. Nuclear export of proteins in plants: AtXPO1 is the export receptor for leucine-rich nuclear export signals in *Arabidopsis thaliana*. *Plant Journal* 20:695-705.
- Haga N, Kato K, Murase M, Araki S, Kubo M, Demura T, Suzuki K, Müller I, Voss U, Jürgens G, Ito M. 2007. R1R2R3-Myb proteins positively regulate cytokinesis through activation of KNOLLE transcription in *Arabidopsis thaliana*. *Development* 134:1101-1110.
- Haga N, Kobayashi K, Suzuki T, Maeo K, Kubo M, Ohtani M, Mitsuda N, Demura T, Nakamura K, Jürgens G, Ito M. 2011. Mutations in MYB3R1 and MYB3R4 cause pleiotropic developmental defects

and preferential down-regulation of multiple G2/M-specific genes in Arabidopsis. *Plant Physiology* 157:706-717.

Haglund K, Dikic I. 2005. Ubiquitylation and cell signalling. *The EMBO Journal* 24:3353–3359.

Halford NG, Paul MJ. 2003. Carbon metabolite sensing and signalling. *Plant Biotechnology Journal* 1:381-398.

Hartmann-Petersen R, Seeger M, Gordon C. 2003. Transferring substrates to the 26S proteasome. *Trends in Biochemical Sciences* 28:26–31.

Hatakeyama S, Yada M, Matsumoto M, Ishida N, Nakayama KI. 2001. U box proteins as a new family of ubiquitin-protein ligases. *Journal of Biological Chemistry* 276:33111-33120.

Hatakeyama S, Nakayama KI. 2003. U-box proteins as a new family of ubiquitin ligases. *Biochemical and Biophysical Research Communications* 302:635-645.

Heidari B, Sayed-Tabatabaei BE, Saeidi G, Kearsey M, Suenaga K. 2011. Mapping QTL for grain yield, yield components, and spike features in a doubled haploid population of bread wheat. *Genome* 54:517-527.

Heine MA, Rankin ML, DiMario PJ. 1993. The Gly/Arg-rich (GAR) domain of *Xenopus* nucleolin facilitates in vitro nucleic acid binding and in vivo nucleolar localization. *Molecular Biology of the Cell* 4:1189-1204.

Hellmann H, Estelle M. 2002. Plant development: regulation by protein degradation. *Science* 297:793–797.

Himanen K, Woloszynska M, Boccardi TM, De Groeve S, Bruno L, Vuylsteke M, Van Lijsebettens M. 2012. Histone H2B monoubiquitination is required to reach maximal transcript levels of circadian clock genes in *Arabidopsis*. *Plant Journal*. doi: 10.1111/j.1365-313X.2012.05071.x.

Hinton JJC. 1955. Resistance of the testa to entry of water into wheat kernel. *Cereal Chemistry* 32:296-306

Hirai S, Oka SI, Adachi E, Kodama H. 2007. The effects of spacer sequences on silencing efficiency of plant RNAi vectors. *Plant Cell Reports* 26:651-659.

Hoecker U, Vasil IK, McCarty DR. 1999. Signaling from the embryo conditions Vp1-mediated repression of alpha-amylase genes in the aleurone of developing maize seeds. *Plant Journal* 19:371-377.

Höfgen R, Willmitzer L. 1988. Storage of competent cells for *Agrobacterium* transformation. *Nucleic Acids Research* 16:9877.

Holzberg S, Brosio P, Gross C, Pogue GP. 2002. Barley stripe mosaic virus-induced gene silencing in a monocot plant. *The Plant Journal* 30:315–327.

Hoppe T, Rape M, Jentsch S. 2001. Membrane-bound transcription factors: regulated release by RIP or RUP. *Current Opinion in Cell Biology* 13:344-348.

Horvath A, Didier A, Koenig J, Exbrayat F, Charmet G, Balfourier F. 2009. Analysis of diversity and linkage disequilibrium along chromosome 3B of bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* 119:1523-1537.

Hoshikawa K. 1962. Studies on the ripening of wheat grain. 4. Influence of temperature upon the development of the endosperm. *Proceedings of the Crop Science Society of Japan* 30: 228-231 (in Japanese with English summary)

Hua Z, Vierstra RD. 2011. The Cullin-RING Ubiquitin-Protein Ligases. *Annual Reviews of Plant Biology* 62:299–334.

Huang XQ, Cöster H, Ganai MW, Röder MS. 2003. Advanced backcross QTL analysis for the identification of quantitative trait loci alleles from wild relatives of wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* 106:1379-1389.

Huibregtse JM, Scheffner M, Beaudenon S, Howley PM. 1995. A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. *Proceedings of National Academy of Sciences of United States of America* 92:5249.

Hurkman WJ, Vensel WH, Tanaka CK, Whitehand L, Altenbach SB. 2009. Effect of high temperature on albumin and globulin accumulation in the endosperm proteome of the developing wheat grain. *Journal of Cereal Science* 49:12–23.

-I-

Ikeda F, Dikic I. 2008. Atypical ubiquitin chains: new molecular signals. *EMBO reports* 9:536-542.

Inzé D, De Veylder L. 2006. Cell cycle regulation in plant development. *Annual Review of Genetic* 40:77-105.

Ito M, Iwase M, Kodama H, Lavis P, Komamine A, Nishihama R, Machida Y, Watanabe A. 1998. A novel cis-acting element in promoters of plant B-type cyclin genes activates M phase-specific transcription. *The Plant Cell* 10:331–341.

Ito M. 2000. Factors controlling cyclin B expression. *Plant Molecular Biology* 43:677–690.

Ito M, Araki S, Matsunaga S, Itoh T, Nishihama R, Machida Y, Doonan JH, Watanabe A. 2001. G2/M-phase-specific transcription during the plant cell cycle is mediated by c-Myb-like transcription factors. *The Plant Cell* 13:1891-1905.

Ito M. 2005. Conservation and diversification of three-repeat Myb transcription factors in plants. *Journal of Plant Research* 118:61-69.

Itahana K, Bhat KP, Jin A, Itahana Y, Hawke D, Kobayashi R, Zhang Y. 2003. Tumor suppressor ARF degrades B23, a nucleolar protein involved in ribosome biogenesis and cell proliferation. *Molecular Cell* 12:1151-1164.

-J-

Jackson PK, Eldridge AG, Freed E, Furstenthal L, Hsu JY, Kaiser BK, Reimann JD. 2000. The lore of the RINGs: substrate recognition and catalysis by ubiquitin ligases. *Trends in Cell Biology* 10:429-439.

Jain M, Li QB, Chourey PS. 2008. Cloning and expression analyses of sucrose non-fermenting-1-related kinase 1 (SnRK1b) gene during development of sorghum and maize endosperm and its implicated role in sugar-to-starch metabolic transition. *Physiologia Plantarum* 134:161-173.

Jang S, Marchal V, Panigrahi KC, Wenkel S, Soppe W, Deng XW, Valverde F, Coupland G. 2008. *Arabidopsis* COP1 shapes the temporal pattern of CO accumulation conferring a photoperiodic flowering response. *EMBO Journal* 27:1277-1288.

Jang IC, Henriques R, Seo HS, Nagatani A, Chua NH. 2010. *Arabidopsis* PHYTOCHROME INTERACTING FACTOR proteins promote phytochrome B polyubiquitination by COP1 E3 ligase in the nucleus. *Plant Cell* 22:2370-2383.

Jeong RD, Chandra-Shekara AC, Barman SR, Navarre D, Klessig DF, Kachroo A, Kachroo P. 2010. Cryptochrome 2 and phototropin 2 regulate resistance protein-mediated viral defense by negatively regulating an E3 ubiquitin ligase. *Proceedings of the National Academy of Sciences of United States of America* 107:13538-13543.

Jimenez-Sanchez G, Childs B, Valle D. 2001. Human disease genes. *Nature* 409:853–855.

Johnsson N, Varshavsky A. 1994. Split ubiquitin as a sensor of protein interactions in vivo. *Proceedings of the National Academy of Science of United States of America* 91:10340-10344.

Jones RA, Qualset CO. 1984. Breeding crops for environmental stress tolerance. In: G.B. Collins and J.G. Petolino, eds. *Application of Genetic Engineering to crop improvement*. p.305-340. Nijhoff/Junk, Dordrecht.

-K-

Kanegae H, Miyoshi K, Hirose T, Tsuchimoto S, Mori M, Nagato Y, Takano M. 2005. Expressions of rice sucrose non-fermenting-1 related protein kinase 1 genes are differently regulated during the caryopsis development. *Plant Physiology and Biochemistry* 43:669-679.

Kang CY, Lian HL, Wang FF, Huang JR, Yang HQ. 2009. Cryptochromes, phytochromes, and COP1 regulate light-controlled stomatal development in *Arabidopsis*. *Plant Cell* 21:2624-2641.

Kassir Y, Rubin-Bejerano I, Mandel-Gutfreund Y. 2006. The *saccharomyces cerevisiae* GSK-3 beta homologs. *Current Drug Targets* 7:1455-1465.

Kato K, Miura H, Sawada S. 2000. Mapping QTLs controlling grain yield and its components on chromosome 5A of wheat. *Theoretical and Applied Genetics* 101:1114–1121.

Keller C, Baggiolini M, 1954: Les stades repères dans la végétation du blé. *Revue Romande D'Agriculture*, Lausanne 10, 17-20.

Khush GS. 2001. Green revolution: the way forward. *Nature Review genetics* 2:815-822.

Kim S, Kang JY, Cho DI, Park JH, Kim SY. 2004. ABF2, an ABRE-binding bZIP factor, is an essential component of glucose signaling and its overexpression affects multiple stress tolerance. *Plant Journal* 40:75-87.

Kobata T, Palta JA, Turner NC. 1992. Rate of development of postanthesis water deficits and grain filling of spring wheat. *Crop Science* 32:1238-1242.

Kosarev P, Mayer KF, Hardtke CS. 2002. Evaluation and classification of RING-finger domains encoded by the Arabidopsis genome. *Genome Biology* 3(4):research0016.1-0016.12.

Kuchel H, Williams KJ, Langridge P, Eagles HA, Jefferies SP. 2007. Genetic dissection of grain yield in bread wheat. I. QTL analysis. *Theoretical and Applied Genetics* 115:1029-1041.

Kumar N, Kulwal PL, Gaur A, Tyagi AK, Khurana JP, Khurana P, Balyan HS, Gupta PK. 2006. QTL analysis for kernel weight in common wheat. *Euphytica*. 151:135-144.

Kumar N, Kulwal PL, Balyan HS, Gupta PK. 2007. QTL mapping for yield and yield contributing traits in two mapping populations of bread wheat. *Molecular Breeding* 19:163-177.

Kurepa J, Walker JM, Smalle J, Gosink MM, Davis SJ, Durham TL, Sung DY, Vierstra RD. 2003. The Small Ubiquitin-like Modifier (SUMO) Protein Modification System in Arabidopsis. Accumulation of SUMO1 and -2 conjugates is increased by stress. *Journal of Biological Chemistry* 278:6862-6872.

Kurepa J, Smalle JA. 2008. Structure, function and regulation of plant proteasomes. *Biochimie* 90:324-335.

Kutschera U, Niklas KJ. 2007. The epidermal-growth-control theory of stem elongation: an old and a new perspective. *Journal of Plant Physiology* 164:1395-1409.

-L-

La Cour T, Kierner L, Mølgaard A, Gupta R, Skriver K, Brunak S. 2004. Analysis and prediction of leucine-rich nuclear export signals. *Protein Engineering, Design & Selection* 17:527-536.

Lam YW, Lamond AI, Mann M, Andersen JS. 2007. Analysis of nucleolar protein dynamics reveals the nuclear degradation of ribosomal proteins. *Current Biology* 17:749-760.

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1:174-181.

Lange A, Mills RE, Lange CJ, Stewart M, Devine SE, Corbett AH. 2007. Classical nuclear localization signals: definition, function, and interaction with importin alpha. *Journal of Biological Chemistry* 282:5101-5105.

Large EC. 1954. Growth stages in cereals illustration of the feekes scale. *Plant Pathology* 3 (4): 128–129.

Lau OS, Deng XW. 2012. The photomorphogenic repressors COP1 and DET1: 20 years later. *Trends in Plant Science* doi:10.1016/j.tplants.2012.05.004

Laudencia-Chingcuanco DL, Stamova BS, You FM, Lazo GR, Beckles DM, Anderson OD. 2007. Transcriptional profiling of wheat caryopsis development using cDNA microarrays. *Plant Molecular Biology* 63:651–668.

Lechner E, Achard P, Vansiri A, Potuschak T, Genschik P. 2006. F-box proteins everywhere. *Current Opinion in Plant Biology* 9:631-638.

- Lee SS, Cho HS, Yoon GM, Ahn JW, Kim HH, Pai HS. 2003. Interaction of NtCDPK1 calcium-dependent protein kinase with NtRpn3 regulatory subunit of the 26S proteasome in *Nicotiana tabacum*. *Plant Journal* 33:825–840.
- Lee J, Zhou P. 2007. DCAFs, the missing link of the CUL4-DDB1 ubiquitin ligase. *Molecular Cell* 26:775-780.
- Lee JH, Yoon HJ, Terzaghi W, Martinez C, Dai M, Li J, Byun MO, Deng XW. 2010. DWA1 and DWA2, two *Arabidopsis* DWD protein components of CUL4-based E3 ligases, act together as negative regulators in ABA signal transduction. *The Plant Cell* 22:1716-1732.
- Li WL, Nelson JC, Chu CY, Shi LH, Huang SH, Liu DJ. 2002. Chromosomal locations and genetic relationships of tiller and spike characters in wheat. *Euphytica* 125:357-366.
- Li X, Qian Q, Fu Z, Wang Y, Xiong G, Zeng D, Wang X, Liu X, Teng S, Hiroshi F, Yuan M, Luo D, Han B, Li J. 2003. Control of tillering in rice. *Nature* 422:618-621.
- Li JR, Zhao W, Li QZ, Ye XG, An BY, Li X, Zhang XS. 2005. RNA silencing of Waxy gene results in low levels of amylose in the seeds of transgenic wheat (*Triticum aestivum* L.). *Acta Genetica Sinica* 32:846–854.
- Li SS, Jia JZ, Wei XY, Zhang XC, Li LZ, Chen HM, Fan YD, Sun HY, Zhao XH, Lei TD et al. 2007. A intervarietal genetic map and QTL analysis for yield traits in wheat. *Molecular Breeding* 20:167-178.
- Li Q, Li L, Yang X, Warburton ML, Bai G, Dai J, Li J, Yan J. 2010a. Relationship, evolutionary fate and function of two maize co-orthologs of rice GW2 associated with kernel size and weight. *BMC Plant Biology* 10:143.
- Li Q, Yang XH, Bai GH, Warburton ML, Mahuku G, Gore M, Dai JR, Li JS, Yan JB. 2010b. Cloning and characterization of a putative GS3 ortholog involved in maize kernel development. *Theoretical and Applied Genetics* 120:753–763.
- Li Y, Fan C, Xing Y, Jiang Y, Luo L, Sun L, Shao D, Xu C, Li X, Xiao J, He Y, Zhang Q. 2011. Natural variation in GS5 plays an important role in regulating grain size and yield in rice. *Nature Genetics* 43:1266-1270.
- Liao BY, Zhang J. 2008. Null mutations in human and mouse orthologs frequently result in different phenotypes. *Proceedings of the National Academy of Sciences of United States of America* 105:6987-6992.
- Lin Q, Wang D, Dong H, Gu S, Cheng Z, Gong J, Qin R, Jiang L, Li G, Wang JL, Wu F, Guo X, Zhang X, Lei C, Wang H, Wan J. 2012. Rice APC/C(TE) controls tillering by mediating the degradation of MONOCULM 1. *Nature Communications* 3:752.
- Liu KJ, Arron JR, Stankunas K, Crabtree GR, Longaker MT. 2007a. Chemical rescue of cleft palate and midline defects in conditional GSK-3 β mice. *Nature* 446:79-82.
- Liu Y, Koornneef M, Soppe WJ. 2007b. The absence of histone H2B monoubiquitination in the *Arabidopsis* hub1 (rdo4) mutant reveals a role for chromatin remodeling in seed dormancy. *Plant Cell* 19:433-444.

Liu LJ, Zhang YC, Li QH, Sang Y, Mao J, Lian HL, Wang L, Yang HQ. 2008. COP1-mediated ubiquitination of CONSTANS is implicated in cryptochrome regulation of flowering in Arabidopsis. *Plant Cell* 20:292-306.

Liu H, Stone SL. 2010. Abscisic acid increases Arabidopsis ABI5 transcription factor levels by promoting KEG E3 ligase self-ubiquitination and proteasomal degradation. *The Plant Cell* 22:2630-2641.

Lizana XC, Riegel R, Gomez LD, Herrera J, Isla A, McQueen-Mason SJ, Calderini DF. 2010. Expansins expression is associated with grain size dynamics in wheat (*Triticum aestivum* L.). *Journal of Experimental Botany* 61:1147-1157.

Lolas IB, Himanen K, Grönlund JT, Lynggaard C, Houben A, Melzer M, Van Lijsebettens M, Grasser KD. 2010. The transcript elongation factor FACT affects Arabidopsis vegetative and reproductive development and genetically interacts with HUB1/2. *The Plant Journal* 61:686-697.

Loukianov A, Yan L, Blechl A, Sanchez A, Dubcovsky J. 2005. Regulation of VRN-1 vernalization genes in normal and transgenic polyploid wheat. *Plant Physiology* 138:2364-2373.

Luo XM, Lin WH, Zhu S, Zhu JY, Sun Y, Fan XY, Cheng M, Hao Y, Oh E, Tian M, Liu L, Zhang M, Xie Q, Chong K, Wang ZY. 2010. Integration of light- and brassinosteroid-signaling pathways by a GATA transcription factor in Arabidopsis. *Developmental Cell* 19:872-883.

-M-

Mao J, Zhang YC, Sang Y, Li QH, Yang HQ. 2005. A role for Arabidopsis cryptochromes and COP1 in the regulation of stomatal opening. *Proceedings of the National Academy of Sciences of United States of America* 102:12270-12275.

Marza F, Bai GH, Carver BF, Zhou WC. 2006. Quantitative trait loci for yield and related traits in the wheat population Ning7840 x Clark. *Theoretical and Applied Genetics* 112:688-698.

Mastrangelo AM, Belloni S, Barilli S, Ruperti B, Di Fonzo N, Stanca AM, Cattivelli L. 2005. Low temperature promotes intron retention in two e-cor genes of durum wheat. *Planta* 221: 705-715.

Mathieu-Rivet E, Gévaudant F, Sicard A, Salar S, Do PT, Mouras A, Fernie AR, Gibon Y, Rothan C, Chevalier C, Hernould M. 2010. Functional analysis of the anaphase promoting complex activator CCS52A highlights the crucial role of endo-reduplication for fruit growth in tomato. *Plant Journal* 62:727-741.

Matsuoka M, Ashikari M. 2007. A quantitative trait locus regulating rice grain width. *Nature Genetics* 39:583-584.

Mattsson K, Pokrovskaja K, Kiss C, Klein G, Szekely L. 2001. Proteins associated with the promyelocytic leukemia gene product (PML)-containing nuclear body move to the nucleolus upon inhibition of proteasome-dependent protein degradation. *Proceedings of the National Academy of Sciences of the United States of America* 98:1012-1017.

Mazzucotelli E, Belloni S, Marone D, De Leonardis A, Guerra D, Di Fonzo N, Cattivelli L, Mastrangelo A. 2006. The e3 ubiquitin ligase gene family in plants: regulation by degradation. *Current Genomics* 7:509-522.

McCartney CA, Somers DJ, Humphreys DG, Lukow O, Ames N, Noll J, Cloutier S, McCallum BD. 2005. Mapping quantitative trait loci controlling agronomic traits in the spring wheat cross RL4452x'AC Domain'. *Genome* 48:870-883.

McGinnis K, Murphy N, Carlson AR, Akula A, Akula C, Basinger H, Carlson M, Hermanson P, Kovacevic N, McGill MA, Seshadri V, Yoyokie J, Cone K, Kaeppler HF, Kaeppler SM, Springer NM. 2007. Assessing the efficiency of RNA interference for maize functional genomics. *Plant Physiology* 143:1441-1451.

McKevith B. 2004. Nutritional aspects of cereals. *Nutrition Bulletin* 29:111-142

Metzger MB, Hristova VA, Weissman AM. 2012. HECT and RING finger families of E3 ubiquitin ligases at a glance. *Journal of Cell Science* 125:531-537.

Miao Y, Zentgraf U. 2010. A HECT E3 ubiquitin ligase negatively regulates Arabidopsis leaf senescence through degradation of the transcription factor WRKY53. *Plant Journal* 63:179-188.

Miki D, Itoh R, Shimamoto K. 2005. RNA silencing of single and multiple members in a gene family in rice. *Plant Physiology* 138:1903-1913.

Misteli T. 2001. Protein dynamics: implications for nuclear architecture and gene expression. *Science* 291:843-847.

Moon J, Estelle M. 2004. The Ubiquitin-proteasome pathway and plant development. *The Plant Cell* 16:3181-3195.

Müller C, Bremer A, Schreiber S, Eichwald S, Calkhoven CF. 2010. Nucleolar retention of a translational C/EBPalpha isoform stimulates rDNA transcription and cell size. *EMBO Journal* 29:897-909.

-N-

Nadaud I, Girousse C, Debiton C, Chambon C, Bouzidi MF, Martre P, Branlard G. 2010. Proteomic and morphological analysis of early stages of wheat grain development. *Proteomics* 10:2901-2910.

Narasimhamoorthy B, Gill BS, Fritz AK, Nelson JC, Brown-Guedira GL. 2006. Advanced backcross QTL analysis of a hard winter wheat x synthetic wheat population. *Theoretical and Applied Genetics* 112:787-796.

Naruoka Y, Talbert LE, Lanning SP, Blake NK, Martin JM, Sherman JD. 2011. Identification of quantitative trait loci for productive tiller number and its relationship to agronomic traits in spring wheat. *Theoretical and Applied Genetics* 123:1043-1053.

Neville M, Stutz F, Lee L, Davis LI, Rosbash M. 1997. The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export. *Current Biology* 7:767-775.

Nguyen HN, Sabelli PA, Larkins BA. 2007. Endoreduplication and programmed cell death in the cereal endosperm. In: Olsen OA, ed, *Endosperm*, Vol 8. Springer-Verlag, Berlin/Heidelberg, pp21-43.

Nicolas ME, Gleadow RM, Dalling MJ. 1984. Effects of drought and high temperature on grain growth in wheat. *Australian Journal of Plant Physiology* 11:553-566.

Ning Y, Jantasuriyarat C, Zhao Q, Zhang H, Chen S, Liu J, Liu L, Tang S, Park CH, Wang X, Liu X, Dai L, Xie Q, Wang GL. 2011a. The SINA E3 ligase OsDIS1 negatively regulates drought response in rice. *Plant Physiology* 157:242-55.

Ning Y, Xie Q, Wang GL. 2011b. OsDIS1-mediated stress response pathway in rice. *Plant Signaling & Behaviour* 6:1684-1686.

-O-

Ohba H, Steward N, Kawasaki S, Berberich T, Ikeda Y, Koizumi N, Kusano T, Sano H. 2000. Diverse response of rice and maize genes encoding homologs of WPK4, an SNF1-related protein kinase from wheat, to light, nutrients, low temperature and cytokinins. *Molecular & General Genetics* 263:359-366.

Olsen OA. 2001. Endosperm development: Cellularization and Cell Fate Specification. *Annual review of plant physiology and plant molecular biology* 52:233–267.

Olsen AN, Ernst HA, Leggio LL, Skriver K. 2005. NAC transcription factors: structurally distinct, functionally diverse. *Trends in Plant Science* 10:79-87.

Oravecz A, Baumann A, Máté Z, Brzezinska A, Molinier J, Oakeley EJ, Adám E, Schäfer E, Nagy F, Ulm R. 2006. CONSTITUTIVELY PHOTOMORPHOGENIC1 is required for the UV-B response in Arabidopsis. *Plant Cell* 18:1975-1990.

Osborne TB. 1924. The vegetables proteins 2nd edition. Longmans, Green and Co Edition. London, England, pp154

Ossareh-Nazari B, Bachelier F, Dargemont C. 1997. Evidence for a role of CRM1 in signal-mediated nuclear protein export. *Science* 278:141-144.

Osterlund MT, Hardtke CS, Wei N, Deng XW. 2000. Targeted destabilization of HY5 during light-regulated development of Arabidopsis. *Nature* 405:462-466.

Oved S, Mosesson Y, Zwang Y, Santonico E, Shtiegman K, Marmor MD, Kochupurakkal BS, Katz M, Lavi S, Cesareni G, Yarden Y. 2006. Conjugation to Nedd8 instigates ubiquitylation and down-regulation of activated receptor tyrosine kinases. *Journal of Biological Chemistry* 281:21640-21651.

-P-

Paolacci AR, Tanzarella OA, Porceddu E, Ciaffi M. 2009. Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. *BMC Molecular Biology* 10:11.

Paux E, Sourdille P, Salse J, Saintenac C, Choulet F, Leroy P, Korol A, Michalak M, Kianian S, Spielmeier W, Lagudah E, Somers D, Kilian A, Alaux M, Vautrin S, Bergès H, Eversole K, Appels R, Safar J, Simkova H, Dolezel J, Bernard M, Feuillet C. 2008. A physical map of the 1-gigabase bread wheat chromosome 3B. *Science* 322:101-4.

Pellegrineschi A, Noguera LM, Skovmand B, Brito RM, Velazquez L, Salgado MM, Hernandez R, Warburton M, Hoisington D. 2002. Identification of highly transformable wheat genotypes for mass production of fertile transgenic plants. *Genome* 45:421-430.

Pendle AF, Clark GP, Boon R, Lewandowska D, Lam YW, Andersen J, Mann M, Lamond AI, Brown JW, Shaw PJ. 2005. Proteomic analysis of the *Arabidopsis* nucleolus suggests novel nucleolar functions. *Molecular Biology of the Cell* 16:260-269.

Persaud A, Rotin D. 2011. Use of proteome arrays to globally identify substrates for E3 ubiquitin ligases. *Methods in Molecular Biology* 759:215-224

Petroski MD, Deshaies RJ. 2005. Function and regulation of cullin-RING ubiquitin ligases. *Nature Reviews Molecular Cell Biology* 6:9-20.

Pfaffl MW. 2004. Quantification Strategies in real-time PCR. In: Bustin, S.A., eds, *The Real-Time PCR Encyclopedia A–Z of Quantitative PCR*. International University Line, 87–120.

Poitrat E. 1999. The potential of liquid biofuels in France. *Renewable Energy* 16, 1084-1089.

Pomeranz Y. 1988. Chemical composition of kernel structures. *Wheat : chemistry and technology*. Volume I, 97-158

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945-959.

Puig O, Caspary F, Rigaut G et al. 2001. The tandem affinity purification (TAP) method: A general procedure of protein complex purification. *Methods* 24:218-229

Puranik S, Sahu PP, Srivastava PS, Prasad M. 2012. NAC proteins: regulation and role in stress tolerance. *Trends in Plant Science* 17:369-381.

-Q-

Quarrie SA, Steed A, Calestani C, Semikhodskii A, Lebreton C, Chinoy C, Steele N, Pljevljakusić D, Waterman E, Weyen J, Schondelmaier J, Habash DZ, Farmer P, Saker L, Clarkson DT, Abugalieva A, Yessimbekova M, Turuspekov Y, Abugalieva S, Tuberosa R, Sanguineti MC, Hollington PA, Aragués R, Royo A, Dodig D. 2005. A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring x SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theoretical and Applied Genetics* 110:865-880.

Qin F, Sakuma Y, Tran LS, Maruyama K, Kidokoro S, Fujita Y, Fujita M, Umezawa T, Sawano Y, Miyazono K, Tanokura M, Shinozaki K, Yamaguchi-Shinozaki K. 2008. *Arabidopsis* DREB2A-interacting proteins function as RING E3 ligases and negatively regulate plant drought stress-responsive gene expression. *The Plant Cell* 20:1693-1707.

-R-

Rabut G, Peter M. 2008. Function and regulation of protein neddylation. 'Protein modifications: beyond the usual suspects' review series. *EMBO Reports* 9:969-976.

Raeker MO, Gaines CS, Finney PL, Donelson T. 1998. Granule size distribution and chemical composition of starches from 12 soft wheat cultivars. *Cereal Chemistry* 75:721-728

Ramya P, Chaubal A, Kulkarni K, Gupta L, Kadoo N, Dhaliwal HS, Chhuneja P, Lagu M, Gupta V. 2010. QTL mapping of 1000-kernel weight, kernel length, and kernel width in bread wheat (*Triticum aestivum* L.). *Journal of Applied Genetics* 51:421-429.

Redman KL, Rechsteiner M. 1989. Identification of the long ubiquitin extension as ribosomal protein S27a. *Nature* 338: 438–440.

Regina A, Bird A, Topping D, Bowden S, Freeman J, Barsby T, Kosar-Hashemi B, Li Z, Rahman S, Morell M. 2006. High-amylose wheat generated by RNA interference improves indices of large-bowel health in rats. *Proceedings of the National Academy of Science of United States of America* 103:3546-3551.

Richards RA. 1988. A tiller inhibitor gene in wheat and its effect on plant growth. *Australian Journal of Agricultural Research* 39:749-757.

Rijven AHGC, Wardlaw IF. 1966. A method for the determination of cell number in plant tissues. *Experimental Cell Research* 41:324-328.

Rizzini L, Favory JJ, Cloix C, Faggionato D, O'Hara A, Kaiserli E, Baumeister R, Schäfer E, Nagy F, Jenkins GI, Ulm R. 2011. Perception of UV-B by the Arabidopsis UVR8 protein. *Science* 332:103-106.

Röder MS, Huang XQ, Börner A. 2008. Fine mapping of the region on wheat chromosome 7D controlling grain weight. *Functional and Integrative Genomics*. 8:79-86.

Rogers SO, Quatrano RS. 1983. Morphological Staging of Wheat Caryopsis Development. *American Journal of Botany* 70:308-311.

Ruden DM, Ma J, Li Y, Wood K, Ptashne M. 1991. Generating yeast transcriptional activators containing no yeast protein sequences. *Nature* 350:250-252.

Ruden DM. 1992. Activating regions of yeast transcription factors must have both acidic and hydrophobic amino acids. *Chromosoma* 101:342-348.

-S-

Sabelli PA, Larkins BA. 2009. The development of endosperm in grasses. *Plant Physiology* 149:14-26

Sabelli PA. 2012. Replicate and die for your own good: Endoreduplication and cell death in the cereal endosperm. *Journal of Cereal Science* 56:9-20.

Saghai-Marooft MA, Soliman KM, Jorgensen RA, Allard RW. 1984. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proceedings of the National Academy of Science of United States of America* 81:8014-8018.

- Saintenac C, Falque M, Martin OC, Paux E, Feuillet C, Sourdille P. 2009. Detailed Recombination Studies Along Chromosome 3B Provide New Insights on Crossover Distribution in Wheat (*Triticum aestivum* L.). *Genetics* 181:393-403.
- Saintenac C, Faure S, Remay A, Choulet F, Ravel C, Paux E, Balfourier F, Feuillet C, Sourdille P. 2011. Variation in crossover rates across a 3-Mb contig of bread wheat (*Triticum aestivum*) reveals the presence of a meiotic recombination hotspot. *Chromosoma* 120:185-198.
- Sambrook J: Molecular cloning : a laboratory manual / Joseph Sambrook, David W. Russell. Cold Spring Harbor, N.Y. :: Cold Spring Harbor Laboratory; 2001.
- Samuel MA, Chong YT, Haasen KE, Aldea-Brydges MG, Stone SL, Goring DR. 2009. Cellular pathways regulating responses to compatible and self-incompatible pollen in Brassica and Arabidopsis stigmas intersect at Exo70A1, a putative component of the exocyst complex. *The Plant Cell* 21:2655-2671.
- San-Segundo PA, Roeder GS. 1999. Pch2 links chromatin silencing to meiotic checkpoint control. *Cell* 97:313-324.
- Santner A, Estelle M. 2010. The ubiquitin-proteasome system regulates plant hormone signaling. *The Plant Journal* 61:1029-1040.
- Sato K, Hayami R, Wu W, Nishikawa T, Nishikawa H, Okuda Y, Ogata H, Fukuda M, Ohta T. 2004. Nucleophosmin/B23 is a candidate substrate for the BRCA1-BARD1 ubiquitin ligase. *Journal of Biological Chemistry* 279:30919-30922.
- Schmidt M, Hanna J, Elsasser S, Finley D. 2005. Proteasome-associated proteins: regulation of a proteolytic machine. *Biological Chemistry* 386:725-737.
- Schmidt-Zachmann MS, Nigg EA. 1993. Protein localization to the nucleolus: a search for targeting domains in nucleolin. *Journal of Cell Science*. 105:799-806.
- Schmittgen TD, Livak KJ. 2008. Analyzing real-time PCR data by the comparative C(T) method. *Nature Protocols* 3:1101-1108.
- Schnyder H, Baum U. 1992. Growth of the grain of wheat (*Triticum aestivum* L.). The relationship between water content and dry matter accumulation. *European Journal of Agronomy* 1:51-57.
- Schubert U, Anton LC, Gibbs J, Norbury CC, Yewdell JW, Binnik JR. 2000. Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature* 404:770-74.
- Seiler C, Harshavardhan VT, Rajesh K, Reddy PS, Strickert M, Rolletschek H, Scholz U, Wobus U, Sreenivasulu N. 2011. ABA biosynthesis and degradation contributing to ABA homeostasis during barley seed development under control and terminal drought-stress conditions. *Journal of Experimental Botany* 62:2615-2632.
- Seo HS, Watanabe E, Tokutomi S, Nagatani A, Chua NH. 2004. Photoreceptor ubiquitination by COP1 E3 ligase desensitizes phytochrome A signaling. *Genes & Development* 18:617-622.
- Shalitin D, Yang H, Mockler TC, Maymon M, Guo H, Whitelam GC, Lin C. 2002. Regulation of Arabidopsis cryptochrome 2 by blue-light-dependent phosphorylation. *Nature* 417:763-767.
- Shah MM, Gill KS, Baenziger PS, Yen Y, Kaeppler SM, Ariyaratne HM. 1999. Molecular Mapping of Loci for Agronomic Traits on Chromosome 3A of Bread Wheat. *Crop Science* 39:1728-1732.

- Shaw P, Brown J. 2012. Nucleoli: composition, function, and dynamics. *Plant Physiology* 158:44-51.
- Shewry PR, Tatham AS, Forde J, Kreis M, Mifflin BJ. 1986. The classification and nomenclature of wheat gluten proteins: a reassessment. *Journal of Cereal Science* 4:97-106.
- Shewry PR. 2009. Wheat. *Journal of Experimental Botany* 60: 1357-1553.
- Shewry PR, Mitchell RAC, Tosi P, Wan Y, Underwood C, Lovegrove A, Freeman J, Toole GA, Mills ENC, Ward JL. 2012. An integrated study of grain development of wheat (cv. Hereward). *Journal of Cereal Science* 56:21-30.
- Singh BK, Jenner CF. 1982. Association between concentrations of organic nutrients in the grain, endosperm cell number and grain dry weight within the ear of wheat. *Australian Journal of Plant Physiology* 9:83-95.
- Sitterlain D. 2005. Quand la virologie dévoile les secrets de l'export des ARN. *Virologie* 9 :191-200.
- Smith NA, Singh SP, Wang MB, Stoutjesdijk PA, Green AG, Waterhouse PM. 2000. Gene expression: Total silencing by intron-spliced hairpin RNAs. *Nature* 407:319-320.
- Snape JW, Foulkes MJ, Simmonds J, Leverington M, Fish LJ, Wang YK, Ciavarrella M. 2007. Dissection gene X environmental effects on wheat yields via QTL and physiological analysis. *Euphytica* 154:401-408.
- Sobhian B, Shao G, Lilli DR, Culhane AC, Moreau LA, Xia B, Livingston DM, Greenberg RA. 2007. RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. *Science* 316:1198-1202.
- Sofield I, Evans LT, Cook MG, Wardlaw IF. 1977a. Factors influencing the rate and duration of grain filling in wheat. *Australian Journal of Plant Physiology* 4:785-797.
- Sofield I, Wardlaw IF, Evans LT, Zee SY. 1977b. Nitrogen, phosphorus and water contents during grain development and maturation in wheat. *Australian Journal of Plant physiology* 4:799-810.
- Soltner D. 1988. Les grandes productions végétales. Collection sciences et techniques agricoles.
- Somers DJ, Isaac P, Edwards K. 2004. A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.) *Theoretical and Applied Genetics* 109:1105-1114.
- Song XJ, Huang W, Shi M, Zhu MZ, Lin HX. 2007. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nature Genetics* 39:623-630.
- Sosulski FW, Imafidon GI. 1990. Amino acid composition and nitrogen-to-protein conversion factors for animal and plant foods. *Journal of Agricultural and Food Chemistry* 38:1351-1356.
- Sourdille P, Cadalen T, Guyomarc'h H, Snape JW, Perretant MR, Charmet G, Boeuf C, Bernard S, Bernard M. 2003. An update of the Courtot x Chinese Spring intervarietal molecular marker linkage map for the QTL detection of agronomic traits in wheat. *Theoretical and Applied Genetics* 106:530-538.
- Spence J, Gali RR, Dittmar G, Sherman F, Karin M, Finley D. 2000. Cell cycle-regulated modification of the ribosome by a variant multiubiquitin chain. *Cell* 102:67-76.

Sreenivasulu N, Altschmied L, Radchuk V, Gubatz S, Wobus U, Weschke W. 2004. Transcript profiles and deduced changes of metabolic pathways in maternal and filial tissues of developing barley grains. *Plant Journal* 37:539-553.

Sreenivasulu N, Radchuk V, Strickert M, Miersch O, Weschke W, Wobus U. 2006. Gene expression patterns reveal tissue-specific signaling networks controlling programmed cell death and ABA-regulated maturation in developing barley seeds. *Plant Journal* 47:310-327.

Sreenivasulu N, Radchuk V, Alawady A, Borisjuk L, Weier D, Staroske N, Fuchs J, Miersch O, Strickert M, Usadel B, Wobus U, Grimm B, Weber H, Weschke W. 2010. De-regulation of abscisic acid contents causes abnormal endosperm development in the barley mutant seg8. *The Plant Journal* 64:589-603.

Sreenivasulu N, Schnurbusch T. 2012. A genetic playground for enhancing grain number in cereals. *Trends in Plant Science* 17:91-101.

Stade K, Ford CS, Guthrie C, Weis K. 1997. Exportin 1 (Crm1p) is an essential nuclear export factor. *Cell* 90:1041-1050.

Stavreva DA, Kawasaki M, Dundr M, Koberna K, Müller WG, Tsujimura-Takahashi T, Komatsu W, Hayano T, Isobe T, Raska I, Misteli T, Takahashi N, McNally JG. 2006. Potential roles for ubiquitin and the proteasome during ribosome biogenesis. *Molecular and Cellular Biology* 26:5131-5145.

Stoddard FL. 1999. Variation in grain mass, grain nitrogen, and starch B-granule content within wheat heads. *Cereal Chemistry* 76:139-144

Stone SL, Anderson EM, Mullen RT, Goring DR. 2003. ARC1 is an E3 ubiquitin ligase and promotes the ubiquitination of proteins during the rejection of self-incompatible Brassica pollen. *The Plant Cell* 15:885-898.

Stone SL, Hauksdóttir H, Troy A, Herschleb J, Kraft E, Callis J. 2005. Functional analysis of the RING-type ubiquitin ligase family of Arabidopsis. *Plant Physiology* 137:13-30.

Stracke R, Werber M, Weisshaar B. 2001. The R2R3-MYB gene family in Arabidopsis thaliana. *Current Opinion in Plant Biology* 4:447-456.

Ström AC, Weis K. 2001. Importin-beta-like nuclear transport receptors. *Genome Biology* 2:REVIEWS3008.

Su Z, Hao C, Wang L, Dong Y, Zhang X. 2011. Identification and development of a functional marker of TaGW2 associated with grain weight in bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* 122:211-223.

Suzuki Y, Iai M, Kamei A, Tanabe Y, Chida S, Yamaguchi S, Zhang Z, Takemoto Y, Shimoizawa N, Kondo N. 2002. Peroxisomal acyl CoA oxidase deficiency. *Journal of Pediatrics* 140:128-130.

Suzuki A, Kogo R, Kawahara K, Sasaki M, Nishio M, Maehama T, Sasaki T, Mimori K, Mori M. 2012. A new PICTURE of nucleolar stress. *Cancer Science* 103:632-637.

Sun C, Berman M, Coward D, Osborne B. 2007. Thickness measurement and crease detection of wheat grains using stereo vision. *Pattern Recognition Letters* 28: 1501-1508.

Sun XY, Wu K, Zhao Y, Kong FM, Han GZ, Jiang HM, Huang XJ, Li RJ, Wang HG, Li SS. 2009. QTL analysis of kernel shape and weight using recombinant inbred lines in wheat. *Euphytica* 165:615-624.

Sun X, Marza F, Ma H, Carver BF, Bai G. 2010. Mapping quantitative trait loci for quality factors in an inter-class cross of US and Chinese wheat. *Theoretical and Applied Genetics* 120:1041-1051.

Surget A, Barron C. 2005. Histologie du grain de blé, *Industrie des Céréales* 145:4-7.

-T-

Taagepera S, McDonald D, Loeb JE, Whitaker LL, McElroy AK, Wang JY, Hope TJ (1998) Nuclear-cytoplasmic shuttling of C-ABL tyrosine kinase. *Proceedings of the National Academy of Sciences of the United States of America* 95:7457–7462.

Tang GQ, Hardin SC, Dewey R, Huber SC. 2003. A novel C-terminal proteolytic processing of cytosolic pyruvate kinase, its phosphorylation and degradation by the proteasome in developing soybean seeds. *Plant Journal* 34:77–93.

Tanaka N, Itoh H, Sentoku N, Kojima M, Sakakibara H, Izawa T, Itoh J, Nagato Y. 2011. The COP1 ortholog PPS regulates the juvenile-adult and vegetative-reproductive phase changes in rice. *Plant Cell* 23:2143-2154.

Tashiro T, Wardlaw IF. 1990. The response to high temperature shock and humidity changes prior to and during the early stages of grain development in wheat. *Australian Journal of Plant Physiology* 17:551-561.

Tateishi K, Omata M, Tanaka K, Chiba T. 2001. The NEDD8 system is essential for cell cycle progression and morphogenetic pathway in mice. *The Journal of Cell Biology* 155:571-579.

Thitisaksakul M, Jiménez RC, Arias MC, Beckles DM. 2012. Effects of environmental factors on cereal starch biosynthesis and composition. *Journal of Cereal Science* 56:67-80.

Tixier MH, Sourdille P, Charmet G, Gay G, Jaby C, Cadalen T, Bernard S, Nicolas P, Bernard M. 1998. Detection of QTLs for crossability in wheat using double-haploid population. *Theoretical and Applied Genetics* 97:1076–1082.

Tosi P, Parker M, Gritsch CS, Carzaniga R, Martin B, Shewry PR. 2009. Trafficking of storage proteins in developing grain of wheat. *Journal of Experimental Botany* 60:979-991.

Tosi P, Gritsch CS, He J, Shewry PR. 2011. Distribution of gluten proteins in bread wheat (*Triticum aestivum*) grain. *Annals of Botany* 108:23-35.

Toyoshima F, Moriguchi T, Wada A, Fukuda M, Nishida E. 1998. Nuclear export of cyclin B1 and its possible role in the DNA damage-induced G2 checkpoint. *EMBO Journal* 17:2728–2735.

Travella S, Klimm TE, Keller B. 2006. RNA interference-based gene silencing as an efficient tool for functional genomics in hexaploid bread wheat. *Plant Physiology* 142:6-20.

Trempe JF. 2011. Reading the ubiquitin postal code. *Current Opinion in Structural Biology* 21:792-801.

-U-

Uauy C, Distelfeld A, Fahima T, Blechl A, Dubcovsky J. 2006. A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. 314:1298-1301.

Ugarte C, Calderini DF, Slafer GA. 2007. Grain weight and grain number responsiveness to pre anthesis temperature in wheat, barley and triticale. *Field Crops Research* 100:240-248.

-V-

Van Lijsebettens M, Grasser KD. 2010. The role of the transcript elongation factors FACT and HUB1 in leaf growth and the induction of flowering. *Plant Signaling & Behavior* 5:715-717.

Van Nocker S, Vierstra RD. 1993. Multiubiquitin chains linked through lysine 48 are abundant in vivo and are competent intermediates in the ubiquitin proteolytic pathway. *Journal of Biological Chemistry* 268:24766–24773.

Vanstraelen M, Baloban M, Da Ines O, Cultrone A, Lammens T, Boudolf V, Brown SC, De Veylder L, Mergaert P, Kondorosi E. 2009. APC/C-CCS52A complexes control meristem maintenance in the *Arabidopsis* root. *Proceedings of the National Academy of Sciences of the United States of America* 106:11806-11811.

Vierstra RD. 2003. The ubiquitin/26S proteasome pathway, the complex last chapter in the life of many plant proteins. *Trends in Plant Science* 8:135–142.

Vierstra RD. 2009. The ubiquitin-26S proteasome system at the nexus of plant biology. *Nature Reviews Molecular Cell Biology* 10:385-397.

Vilhar B, Kladnik A, Blejec A, Chourey PS, Dermastia M. 2002. Cytometrical evidence that the loss of seed weight in the miniature1 seed mutant of maize is associated with reduced mitotic activity in the developing endosperm. *Plant Physiology* 129:23-30.

-W-

Wada A, Fukuda M, Mishima M, Nishida E. 1998. Nuclear export of actin: a novel mechanism regulating the subcellular localization of a major cytoskeletal protein. *EMBO Journal* 17:1635–1641.

Wagstaff KM, Jans DA. 2009. Importins and beyond: Non-conventional nuclear transport mechanisms. *Traffic* 10:1188-1198.

Wan Y, Poole RL, Huttly AK, Toscano-Underwood C, Feeney K, Welham S, Gooding MJ, Mills C, Edwards KJ, Shewry PR, Mitchell RA. 2008. Transcriptome analysis of grain development in hexaploid wheat. *BMC Genomics* 9:121.

Wang E, Wang J, Zhu X, Hao W, Wang L, Li Q, Zhang L, He W, Lu B, Lin H, Ma H, Zhang G, He Z. 2008. Control of rice grain-filling and yield by a gene with a potential signature of domestication. *Nature Genetics* 40:1370-1374.

- Wang RX, Hai L, Zhang XY, You GX, Yan CS, Xiao SH. 2009. QTL mapping for grain filling rate and yield-related traits in RILs of the Chinese winter wheat population Heshangmai x Yu8679. *Theoretical and Applied Genetics* 118:313-325.
- Wang L, Ge H, Hao C, Dong Y, Zhang X. 2012. Identifying loci influencing 1,000-kernel weight in wheat by microsatellite screening for evidence of selection during breeding. *PLoS One* 7:e29432.
- Wardlaw IF, Moncur L. 1995. The response of wheat to high temperature following anthesis. I. The rate and duration of kernel filling. *Australian Journal of Plant Physiology* 22:391-397.
- Weissman AM. 2001. Themes and variations on ubiquitylation. *Nature Reviews Molecular Cell Biology* 2:169-178.
- Wendt KS, Vodermaier HC, Jacob U, Gieffers C, Gmachl M, Peters JM, Huber R, Sondermann P. 2001. Crystal structure of the APC10/DOC1 subunit of the human anaphase-promoting complex. *Nature Structural Biology* 8:784-788.
- Weng J, Gu S, Wan X, Gao H, Guo T, Su N, Lei C, Zhang X, Cheng Z, Guo X, Wang J, Jiang L, Zhai H, Wan J. 2008. Isolation and initial characterization of GW5, a major QTL associated with rice grain width and weight. *Cell Research* 18:1199-1209.
- Wesley SV, Helliwell CA, Smith NA, Wang MB, Rouse DT, Liu Q, Gooding PS, Singh SP, Abbott D, Stoutjesdijk PA, Robinson SP, Gleave AP, Green AG, Waterhouse PM. 2001. Construct design for efficient, effective and high-throughput gene silencing in plants. *The Plant Journal* 27:581-590.
- Wiegand CL, Cuellar JA. 1981. Duration of grain filling and kernel weight of wheat as affected by temperature. *Crop Science* 21:95-101.
- Wilkinson S, Kudoyarova GR, Veselov DS, Arkhipova TN, Davies WJ. 2012. Plant hormone interactions: innovative targets for crop breeding and management. *Journal of Experimental Botany* 63:3499-3509.
- Woo HR, Chung KM, Park JH, Oh SA, Ahn T, Hong SH, Jang SK, Nam HG. 2001. ORE9, an F-box protein that regulates leaf senescence in Arabidopsis. *The Plant Cell* 13:1779-1790.
- Woodger FJ, Jacobsen JV, Gubler F. 2004. GMPOZ, a BTB/POZ domain nuclear protein, is a regulator of hormone responsive gene expression in barley aleurone. *Plant and Cell Physiology* 45:945-950.
- Wright M, Dawson J, Dunder E, Suttie J, Reed J, Kramer C, Chang Y, Novitzky R, Wang H, Artim-Moore L. 2001. Efficient biolistic transformation of maize (*Zea mays* L.) and wheat (*Triticum aestivum* L.) using the phosphomannose isomerase gene, *pmi*, as the selectable marker. *Plant Cell Reports* 20:429-436.
- Wu W, Koike A, Takeshita T, Ohta T. 2008. The ubiquitin E3 ligase activity of BRCA1 and its biological functions. *Cell Division* 3:1.
- Wu Y, Zhang D, Chu JY, Boyle P, Wang Y, Brindle ID, De Luca V, Després C. 2012. The Arabidopsis NPR1 protein is a receptor for the plant defense hormone salicylic acid. *Cell reports* 1:639-647.
- Wu-Baer F, Lagrazon K, Yuan W, Baer R. 2003. The BRCA1/BARD1 heterodimer assembles polyubiquitin chains through an unconventional linkage involving lysine residue K6 of ubiquitin. *Journal of Biological Chemistry* 278:34743-34746.

-X-

- Xia Z, Liu Q, Wu J, Ding J. 2012. ZmRFP1, the putative ortholog of SDIR1, encodes a RING-H2 E3 ubiquitin ligase and responds to drought stress in an ABA-dependent manner in maize. *Gene* 495:146–153.
- Xie DX, Feys BF, James S, Nieto-Rostro M, Turner JG. 1998. COI1: an Arabidopsis gene required for jasmonate-regulated defense and fertility. *Science* 280:1091–1094.
- Xie Q, Guo HS, Dallman G, Fang S, Weissman AM, Chua NH. 2002. SINAT5 promotes ubiquitin-related degradation of NAC1 to attenuate auxin signals. *Nature* 419:167-170.
- Xing Y and Zhang Q. 2010. Genetic and molecular bases of rice yield. *Annual Review of Plant Biology* 61:421-442.
- Xirodimas DP, Saville MK, Bourdon JC, Hay RT, Lane DP. 2004. Mdm2-mediated NEDD8 conjugation of p53 inhibits its transcriptional activity. *Cell* 118:83-97.
- Xirodimas DP, Sundqvist A, Nakamura A, Shen L, Botting C, Hay RT. 2008. Ribosomal proteins are targets for the NEDD8 pathway. *EMBO Reports* 9:280-286.
- Xu G, Zhang J, Lam HM, Wang Z, Yang J. 2007. Hormonal changes are related to the poor grain filling in the inferior spikelets of rice cultivated under non-flooded and mulched condition. *Field Crops Research* 101:53-61.
- Xu C, Wang Y, Yu Y, Duan J, Liao Z, Xiong G, Meng X, Liu G, Qian Q, Li J. 2012. Degradation of MONOCULM 1 by APC/CTAD1 regulates rice tillering. *Nature Communications* 3:750.

-Y-

- Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J. 2000. Positional cloning of the wheat vernalization gene VRN1. *Proceedings of the National Academy of Sciences of the United States of America* 100:6263-6268.
- Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, SanMiguel P, Bennetzen JL, Echenique V, Dubcovsky J. 2004. The wheat VRN2 gene is a flowering repressor down-regulated by vernalization. *Science* 303:1640-1644.
- Yan S, Zou G, Li S, Wang H, Liu H, Zhai G, Guo P, Song H, Yan C, Tao Y. 2011. Seed size is determined by the combinations of the genes controlling different seed characteristics in rice. *Theoretical and Applied Genetics* 123:1173-1181.
- Yang J, Zhang J, Wang Z, Zhu Q, Wang W. 2001a. Hormonal changes in the grains of rice subjected to water stress during grain filling. *Plant Physiology* 127:315-323.
- Yang Y, Kuang Y, Montes De Oca R, Hays T, Moreau L, Lu N, Seed B, D'Andrea AD. 2001b. Targeted disruption of the murine Fanconi anemia gene, *Fancg/Xrcc9*. *Blood* 98:3435–3440.
- Yang P, Fu H, Walker J, Papa CM, Smalle J, Ju YM, Vierstra RD. 2004. Purification of the Arabidopsis 26S proteasome: biochemical and molecular analyses revealed the presence of multiple isoforms. *Journal of Biological Chemistry* 279:6401–6413.

Yang Z, Bai Z, Li X, Wang P, Wu Q, Yang L, Li L, Li X. 2012. SNP identification and allelic-specific PCR markers development for TaGW2, a gene linked to wheat kernel weight. *Theoretical and Applied Genetics* doi: 10.1007/s00122-012-1895-6.

Yanhui C, Xiaoyuan Y, Kun H, Meihua L, Jigang L, Zhaofeng G, Zhiqiang L, Yunfei Z, Xiaoxiao W, Xiaoming Q, Yunping S, Li Z, Xiaohui D, Jingchu L, Xing-Wang D, Zhangliang C, Hongya G, Li-Jia Q. 2006. The MYB transcription factor superfamily of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Molecular Biology* 60:107-124.

Yee D, Goring DR. 2009. The diversity of plant U-box E3 ubiquitin ligases: from upstream activators to downstream target substrates. *Journal of Experimental Botany* 60:1109-1121.

Yi C, Deng XW. 2005. COP1-from plant photomorphogenesis to mammalian tumorigenesis. *Trends in Cell Biology* 15:618-625.

Yoshida S, Asakawa K, Toh-e A. 2002. Mitotic exit network controls the localization of Cdc14 to the spindle pole body in *Saccharomyces cerevisiae*. *Current Biology* 12: 944–950.

Young TE, Gallie DR. 2000. Programmed cell death during endosperm development. *Plant Molecular Biology* 44:283-301.

Yu JW, Rubio V, Lee NY, Bai S, Lee SY, Kim SS, Liu L, Zhang Y, Irigoyen ML, Sullivan JA, Zhang Y, Lee I, Xie Q, Paek NC, Deng XW. 2008. COP1 and ELF3 control circadian function and photoperiodic flowering by regulating GI stability. *Molecular Cell* 32:617-630.

Yue SJ, Li H, Li YW, Zhu YF, Guo JK, Liu YJ, Chen Y, Jia X. 2007. Generation of transgenic wheat lines with altered expression levels of 1Dx5 high-molecular-weight glutenin subunit by RNA interference. *Journal of Cereal Science* 47:153-161.

-Z-

Zadoks JC, Chang TT, Konzak CF. 1974. A decimal code for the growth stages of cereals. *Weed Research* 14:415-421.

Zee SY, O'Brien TP. 1970. Studies on the ontogeny of the pigment strand in the caryopsis of wheat. *Australian Journal of biological Science* 23:1153-1171.

Zee SY. 1975. Accumulation of an adcrusting substance in the pigment strand in the caryopsis of wheat. *Canadian Journal of Botany* 53:2246-2250.

Zeng LR, Park CH, Venu RC, Gough J, Wang GL. 2008. Classification, expression pattern, and E3 ligase activity assay of rice U-box-containing proteins. *Molecular Plant* 1:800-815.

Zhang YY, Yang CW, Li Y, Zheng NY, Chen H, Zhao QZ, Gao T, Guo HS, Xie Q. 2007. SDIR1 is a RING finger E3 ligase that positively regulates stress-responsive abscisic acid signaling in Arabidopsis. *The Plant Cell* 19:1912–1929.

Zhang LY, Liu DC, Guo XL, Yang WL, Sun JZ, Wang DW, Zhang A. 2010. Genomic distribution of quantitative trait loci for yield and yield-related traits in common wheat. *Journal of integrative Plant Biology* 52:996-1007.

Zhang L, Zhao G, Jia J, Liu X, Kong X. 2012. Molecular characterization of 60 isolated wheat MYB genes and analysis of their expression during abiotic stress. *Journal of Experimental Botany* 63:203-214.

Zheng N, Wang P, Jeffrey PD, Pavletich NP. 2000. Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. *Cell* 102:533-539.

Zheng J, Chen F, Wang Z, Cao H, Li X, Deng X, Soppe WJJ, Li Y, Liu Y. 2012. A novel role for histone methyltransferase KYP/SUVH4 in the control of Arabidopsis primary seed dormancy. *New Phytologist* 193: 605–616.

Zhu G, Ye N, Yang J, Peng X, Zhang J. 2011. Regulation of expression of starch synthesis genes by ethylene and ABA in relation to the development of rice inferior and superior spikelets. *Journal of Experimental Botany* 62:3907-39016.

Zohary D, Hopf M. 2000. Domestication of plants in the old world: the origin and spread of cultivated plants in West Asia, Europe, and the Nile Valley. Oxford University Press.

Article Scientifique:

Down-regulation of the *TaGW2* gene by RNA interference results in decreased grain size and weight in wheat

Julie Bednarek, Aurélia Boulaflois, Christine Girousse, Catherine Ravel, Caroline Tassy,
Pierre Barret, Mohamed Fouad Bouzidi, Said Mouzeyar

Journal of Experimental Botany, 2012

RESEARCH PAPER

Down-regulation of the *TaGW2* gene by RNA interference results in decreased grain size and weight in wheat

Julie Bednarek¹, Aurélia Boulaflous¹, Christine Girousse², Catherine Ravel², Caroline Tassy², Pierre Barret², Mohamed Fouad Bouzidi¹ and Said Mouzeyar^{1,*}

¹ Université Blaise Pascal, UMR 1095 GDEC, 24 avenue des Landais, F-63177 Aubière, France

² INRA, UMR 1095 GDEC, 234 avenue du Brézat, F-63100 Clermont-Ferrand, France

* To whom correspondence should be addressed. E mail: said.mouzeyar@univ-bpclermont.fr

Received 17 April 2012; Revised 6 July 2012; Accepted 6 August 2012

Abstract

For important food crops such as wheat and rice, grain yield depends on grain number and size. In rice (*Oryza sativa*), *GW2* was isolated from a major quantitative trait locus for yield and encodes an E3 RING ligase that negatively regulates grain size. Wheat (*Triticum aestivum*) has *TaGW2* homologues in the A, B, and D genomes, and polymorphisms in *TaGW2-A* were associated with grain width. Here, to investigate *TaGW2* function, RNA interference (RNAi) was used to down-regulate *TaGW2* transcript levels. Transgenic wheat lines showed significantly decreased grain size-related dimensions compared with controls. Furthermore, *TaGW2* knockdown also caused a significant reduction in endosperm cell number. These results indicate that *TaGW2* regulates grain size in wheat, possibly by controlling endosperm cell number. Wheat and rice *GW2* genes thus seem to have divergent functions, with rice *GW2* negatively regulating grain size and *TaGW2* positively regulating grain size. Analysis of transcription of *TaGW2* homologues in developing grains suggested that *TaGW2-A* and *-D* act in both the division and late grain-filling phases. Furthermore, biochemical and molecular analyses revealed that *TaGW2-A* is a functional E3 RING ubiquitin ligase with nucleocytoplasmic subcellular partitioning. A functional nuclear export sequence responsible for *TaGW2-A* export from the nucleus to the cytosol and retention in the nucleolus was identified. Therefore, these results show that *TaGW2* acts in the regulation of grain size and may provide an important tool for enhancement of grain yield.

Key words: Divergent function, E3 RING ligase, grain size, orthologous, RNA interference, *TaGW2*, wheat.

Introduction

Wheat is one of the world's major cereal crops and is of considerable importance to human nutrition, supplying one-fifth of the calories consumed by humans. Wheat is a hexaploid plant, having three homoeologous genomes, A, B, and D. The wheat grain is a caryopsis and consists of the embryo, the endosperm, and the teguments, with the endosperm being the primary storage site for starch and protein. Wheat grain development can generally be divided into five successive but overlapping stages based on endosperm development (Simmonds and O'Brien, 1981): fertilization [0 days post-anthesis (DPA)], coenocytic endosperm, cellularization, and cell division (these three processes occurring from 1 to 5 DPA), early grain filling (6–13

DPA), late grain filling (14–24 DPA), and grain desiccation (25–38 DPA).

In cereals, grain yield is a complex agronomic trait that is controlled by quantitative trait loci (QTLs) and affected by environmental factors. It can be broken down into a number of components that are also under QTL control, such as spike number per plant, grain weight per spike, spikelet number per spike, and 1000 grain weight (TGW). Grain weight is the most important component of grain yield and is largely determined by grain size, which is specified using its three dimensions (length, width, and thickness), and the degree of filling. In wheat, several QTLs for grain yield have been identified (Kato *et al.*, 2000;

McCartney et al., 2005; Röder et al., 2008). However, no gene controlling a QTL for wheat grain size has yet been isolated and fully characterized. In contrast, in rice, several genes controlling QTLs for yield component traits have been cloned and functionally studied, such as *GS3* (Fan et al., 2006), *GW2* (Song et al., 2007), and, more, recently *GW5* (Weng et al., 2008).

GW2 was isolated from a QTL controlling grain width and weight. *GW2* was mapped on the short arm of rice chromosome 2 and the 1278 bp cDNA encodes a RING-type E3 ligase. E3 ligases (E3s) mediate specifically ubiquitination on protein substrates in the ubiquitin–26S proteasome system (Vierstra, 2009). The enzymatic reaction is carried out by a set of three enzymes, E1, E2, and E3, and the latter catalyses the transfer of ubiquitin from E2 to the targeted protein. The *GW2* protein is a 425 residue polypeptide of ~47 kDa, with a C5HC2 RING domain in its N-terminal region (Song et al., 2007). Loss of *GW2* function increases the number of cells in the spikelet hull, yielding wider spikelet hulls and consequently wider rice grains, and also accelerates the grain milk-filling rate, resulting in heavier grains. The *GW2* protein therefore functions as a negative regulator of cell division. The substrate targeted by *GW2* during cell division in the rice spikelet hull is still unknown and the mechanism by which *GW2* alters the grain milk-filling rate also remains unclear. Recently, *GW2* homologues were identified in maize and wheat (Li et al., 2010a; Su et al., 2011). In maize, *ZmGW2-CHR4* is significantly associated with grain width and 100 grain weight (HGW). Moreover, *ZmGW2-CHR4* was located in a HGW QTL (Li et al., 2010a). In bread wheat, *TaGW2* is constitutively expressed, and *TaGW2* homologues in the A, B, and D genomes are located on the homoeologous group 6 chromosomes (Su et al., 2011). *TaGW2-A* was mapped on the short arm of chromosome 6A, near the centromere. Association analyses in a Chinese common wheat core collection revealed that a marker in the promoter region of *TaGW2-A* was significantly associated with grain width and TGW. This marker was also found to be strongly associated with heading and maturation date (Su et al., 2011). More recently, Yang and colleagues (2012) identified a single T base insertion in the eighth exon of *TaGW2-A* leading to a premature stop codon. These authors found this single nucleotide polymorphism (SNP) genetically associated with grain width in an F_2 segregating population (Lankaodali×Chinese Spring).

The aim of this study was to investigate the function of *TaGW2* in wheat grain development. The hexaploid nature of the wheat genome makes finding and combining mutations in genes that are expressed by the three genomes a challenging proposition. In contrast, gene silencing through RNA interference (RNAi) can simultaneously suppress target gene expression from each locus. Here, RNAi was used to silence the three expressed copies of *TaGW2*, and the results showed that the down-regulation of *TaGW2* expression resulted in a significant reduction in final grain weight and size. To gain further insights into *TaGW2* function, biochemical and molecular analyses were performed. First, the results showed that *TaGW2-A* and *-D* are mainly expressed during the division and late grain-filling stages. In addition, the *TaGW2-A* protein is a functional E3 RING ligase, like its rice counterpart, but displays nucleocytoplasmic localization. An N-terminal LxxLxL-type nuclear

export signal (NES) that is responsible for shuttling *TaGW2-A* between the nucleus and cytosol and retaining *TaGW2-A* in the nucleolus was identified.

Materials and methods

Plant material and growth conditions

Seeds from *Triticum aestivum* cv Récital and transgenic plants were grown as described in Capron et al. (2012) and in the Supplementary information SI available at JXB online.

For the determination of grain developmental stages, the first flowering spikelets of all tillers on each plant were tagged at anthesis and only grains from florets with the same anthesis date were harvested. The developmental stages of grain were evaluated in terms of growth degree days ($^{\circ}\text{Cd}$) from anthesis; that is, the sum of daily mean temperatures accumulated over the period considered.

For transgenic experiments, immature seeds were harvested 12–14 DAP and sterilized, and immature embryos were isolated for bombardment.

Generation of RNAi lines

The RNAi construct (pTaGW2-RNAi) was made in the vector pSTAR-GATE supplied by CSIRO (<http://www.pi.csiro.au/rnai/vectors.htm>).

RNAi lines were produced using particle bombardment-mediated transformation of immature cv Récital embryos. The pTaGW2-RNAi plasmid was digested with *NotI*, and the fragment containing the RNAi cassette was co-bombarded with a phospho-mannose isomerase selection cassette with a 2:1 molar ratio. Regeneration and selection of the transformed plants were performed essentially as described by Wright et al. (2001) and Pellegrineschi et al. (2002).

Genotyping of transgenic plants was performed by PCR of genomic DNA using PDK–intron primer pairs (Supplementary Table S1). In positive and negative transgenic lines, endogenous transcription levels of each *TaGW2* were investigated by quantitative real-time PCR (qRT-PCR) as mentioned below, in RNA samples from leaves, using the RNAiTaGW2endo primer pairs (Supplementary Table S1 at JXB online).

Quantitative real-time PCR

Total RNA and first-strand cDNA were obtained as described by Capron et al. (2012) and in Supplementary information SI at JXB online. Absence of genomic DNA contamination was checked by PCR using the PDI–intron primer pair (Paolacci et al., 2009).

Specificity and PCR efficiency of primer pairs were determined as described in Supplementary information SI. Only pairs with efficiencies between 85% and 100% were used for quantification. qRT-PCR assays were performed on 4 μl of cDNA diluted 1:40. Three independent biological replicates were used for quantification analyses and three technical replicates were analysed per biological replicate. PCRs were cycled as described in Capron et al. (2012). Relative expression was determined using the ΔCt method corrected for primer efficiencies (Pfaffl, 2004). Data were normalized to the expression of the *Ta2776* gene (Paolacci et al., 2009). To compare *TaGW2* homoeologue expression, the relative expression values of *TaGW2-A* were calculated using the $2^{-\Delta\Delta\text{Ct}}$ method, with the 40 $^{\circ}\text{Cd}$ stage as a reference sample for $\Delta\Delta\text{Ct}$, according to the method described by Schmittgen and Livak (2008).

In RNAi lines, two independent technical replicates were used for quantification analyses and three replicates were analysed per technical replicate. The *Ta54227* gene was used as the internal control. Relative expression values were calculated using the $2^{-\Delta\Delta\text{Ct}}$ method, with the mean of ΔCt values of the null segregants from the transgenic event as a calibrator.

Phenotyping of RNAi transgenic lines

Each independent transgenic event was conducted and phenotyped, separately in time. Phenotyping of T_2 progeny from T105 and T111 events

was conducted between June 2010 and January 2011 and between November 2010 and May 2011, respectively.

For each transgenic event, spikes from all tillers of each plant were tagged at anthesis when the first spikelets of the spike flowered, and the date was recorded. For both transgenic events, spikes were harvested at full maturity (1000 °Cd after anthesis). The numbers of spikelets per spike, and the numbers of grain per spike, were then determined. Measures of individual grain fresh mass, grain dry mass, volume, water mass estimated as the difference between fresh and dry masses, length, width, and thickness were carried out on the two basal grains of the two central spikelets on each spike of the plant. Grain volume was estimated using a water displacement method (pycnometry), and grain size traits were determined using a scaled camera-assisted system. Grain fresh mass, volume, length, width, and thickness were obtained from freshly harvested caryopses and grain dry mass after oven drying of the grains at 80 °C for 48 h. All remaining grains from a spike were counted, pooled, and weighed, and ground to whole meal flour using a 6800 Cyclotec mill. Total grain nitrogen content was then determined with the Dumas combustion method (AOAC method no. 992.23) using a FlashEA 1112 N/Protein Analyzer (Thermo Electron). Each extract was analysed in duplicate. Total grain protein content per grain was calculated by multiplying the grain nitrogen content by 5.7 (Sosulski and Imafidon, 1990).

For the T105 transgenic event, basal grains of central spikelets of spikes were also harvested at ~260 °Cd, and the number of cells in the endosperm was then counted according to the method described by Singh and Jenner (1982).

In vitro autoubiquitination assay, Agrobacterium-mediated transient expression in Nicotiana tabacum, and confocal laser scanning microscopy (CLSM)

In vitro ubiquitination assays were carried out as described by Song *et al.* (2007), with slight modifications (Supplementary information S1 at JXB online). *Nicotiana tabacum* leaf epidermal cells were transiently co-transformed via agroinfiltration as described by Boulaflos *et al.* (2009). CLSM analyses were performed as described in Supplementary information S1.

Statistical analysis

For each transgenic event, variance analysis was performed using SAS/STAT® software (GLM procedure, SAS Institute, 2002). The comparisons between means were performed using a Student–Newman–Keuls (SNK) test (at the 5% level).

In RNAi experiments and whole grain transcriptional analysis, significant differences were calculated by analysis of variance (ANOVA) using Scheffé's test. A value of $P < 0.01$ was taken as the criterion of significance.

Results

Cloning and characterization of *TaGW2* homoeologous cDNAs

To isolate the homoeologous cDNAs encoding *TaGW2*, primers (*TaGW2* FL CDS primer pair, Supplementary Table S1 at JXB online) were designed from the rice *GW2* CDS sequence (EF447275.1), and then cDNAs encoding *TaGW2* from leaves of *T. aestivum* were amplified and cloned (cv Récltal). Sequence analysis revealed two different cDNAs. Two groups of wheat expressed sequence tags (ESTs) were also found in GenBank, one group corresponding to each of the cDNAs, and a single EST (CK206178) with slight differences from the other two groups. Each group of sequences was then assigned to the short arm of chromosomes 6A, 6B, and 6D by PCR on

aneuploid lines of *T. aestivum* cv Chinese Spring (Sears, 1966), using copy-specific primers (*TaGW2*-6A, -6B, and -6D primer pairs, designed from an alignment of the cDNA sequences; Supplementary Table S1 and Fig. S1). These results revealed that the full-length sequences of *TaGW2*-A and -B (GenBank accession numbers JN896622 and JN896623, respectively) were thus obtained. The EST CK206178 corresponded to a 3'-partial sequence of *TaGW2*-D (from 679 bp to 1275 bp of the *TaGW2*-A and -B cDNAs, Supplementary Fig. S1).

Production of RNAi lines

To investigate the function of *TaGW2* during wheat grain development, RNAi was performed to reduce the transcript levels of all *TaGW2* homoeologues. The RNAi construct was made with a 1242 bp fragment of the *TaGW2*-A cDNA (nucleotides 34–1275). The 1242 bp sequence was used to query the wheat genomic databases available for sequences that could be potentially targeted by the RNAi construct. Blastn searches returned 13 and eight hits with stretches of identity of ≥ 20 bp in CerealsDB (www.cerealsdb.uk.net) and URGI (<http://urgi.versailles.inra.fr>) databases, respectively. Four hits with no homology with known genes were found in URGI. One sequence on each of the chromosomes 7AL and 3AL and two sequences on the 2AS chromosome were found together to correspond to non-coding DNA or to a Gypsy-like retrotransposon in the provisional wheat genome sequence. All remaining hits were considered to be *TaGW2* genes.

Two series of 354 and 246 immature embryos of winter hexaploid wheat variety Récltal were transformed by co-bombardment with the RNAi construct and a selection cassette. Two independent transgenic plants, T105 and T111, were identified. Transgenic T₀ plants were self-pollinated and the presence or absence of the transgene in T₁ progeny was determined by PCR on genomic DNA. QRT-PCR analysis in T₁ progeny enabled the identification of one positive T105 plant with a substantial reduction of overall *TaGW2* expression (75% reduction) and four positive T111 plants with an average reduction in overall *TaGW2* expression of 50%. The plants were self-pollinated and the resulting T₂ progeny were genotyped. QRT-PCR of T₂ progeny identified 21 and eight plants from T105 and T111 lines, respectively, with an average reduction in overall *TaGW2* expression of 40% compared with null-segregant control lines (the overall *TaGW2* reduction level ranged from 20% to 76% and from 35% to 64% in positive plants from T105 and T111 lines, respectively, Fig. 1A). The expression levels of the *TaGW2* genes in the RNAi lines were analysed using homoeologue-specific primers (*TaGW2*-A, -B, and -D reverse primers; Supplementary Table S1 at JXB online). In both transgenic lines, a significant reduction in transcript levels of each *TaGW2* homoeologue was observed (Fig. 1A). The A and B copies of *TaGW2* were reduced by almost 45% in both transgenic events compared with control lines, and the D copy by 39% and 22% in T105 and T111 transgenic lines, respectively.

Measurement of grain yield-related traits in RNAi lines

To determine whether reduction of *TaGW2* affected grain yield, T₂ transgenic and null-segregant plants from T105 and T111 lines

were phenotyped at maturity for grain yield and size-related traits. Significant effects on grain size were observed in mature grains of both transgenic lines compared with the controls (Fig. 1B–D). Spikelet position on the spike impacts grain weight and nitrogen accumulation (Calderini and Ortiz-Monasterio, 2003; Andersson et al., 2004). Therefore, to compare traits accurately between transgenic and control lines, all measurements were performed on the two basal grains of the two central spikelets from the main spike of each plant. Individual grain fresh mass was reduced in transgenic plants compared with controls by 38.6% and 22.2%, grain dry mass was reduced by 37.3% and 21.9%, and grain water mass by 31% and 25%, in T105 and T111 lines, respectively (Fig. 1E). Transgenic plants also displayed a 34.6% and 20.4% reduction in grain volume, a 10% and 11.8% reduction in grain width, and a 13.3% and 9.1% reduction in thickness, in T105 and T111 lines, respectively. No significant differences were observed in either transgenic line in the number of grains per spike, the number of spikelets per spike, the anthesis date, or the nitrogen and protein contents per grain. Differences in grain length and number of spikes per plant in transgenic plants compared with the controls were only observed in the T111 line (Fig. 1E).

In T₂ progeny from the T105 line, the number of cells in the endosperm of grains harvested at 260 °Cd (considered as the end of the division stage in cv Récltal; Nadaud et al., 2010) was also phenotyped. Cell numbers in the endosperm were significantly reduced by ~25% in transgenic plants compared with control plants (Fig. 1E). Unfortunately, endosperm cell numbers of T₂ progeny from the T111 line could not be counted because of the limited number of spikes yielded by these plants (transgenic and null-segregant control, Fig. 1E). Taken together, these results suggest that *TaGW2* is a positive regulator of wheat grain size and grain weight, unlike *GW2* in rice, and that *TaGW2* is a positive regulator of cell number in the endosperm.

Expression of *TaGW2* homoeologues during grain development

GW2 and its homologues are constitutively expressed in rice, maize, and wheat (Song et al., 2007; Li et al., 2010a; Su et al., 2011), but these studies only covered grain division stages, and the specific expression patterns of each *TaGW2* homoeologue were not investigated. To analyse *TaGW2* homoeologue expression during grain development, a qRT-PCR assay was performed using homoeologue-specific primer pairs (Supplementary Table S1 at JXB online), on eight grain developmental stages, at times measured in cumulative degree days (°Cd; see Materials and methods), in *T. aestivum* cv Récltal from 40 °Cd to 800 °Cd. All transcripts were detected from 40 °Cd (Fig. 2). *TaGW2-A* and *-D* were mainly expressed during the beginning of the division stage (40–120 °Cd) and the late grain-filling stage (600 °Cd). A high level of *TaGW2-D* transcript was maintained during grain desiccation (800 °Cd). *TaGW2-B* transcript mainly accumulated during the late grain-filling stage (600 °Cd). Therefore, the expression patterns of *TaGW2-A* and *-D* strongly suggested their involvement in both the division and late grain-filling phases of wheat grain development. Furthermore, the high level of *TaGW2-D* at 800 °Cd suggested that it may have a role in grain

maturation. *TaGW2-B* may only act in late grain filling. Based on these results, the A and D copies appeared to have major roles in the regulation of division and filling in wheat grain development. Likewise, Su and colleagues (2011) reported the A-copy to be associated with grain width. With this in mind, it would thus be interesting to characterize further the *TaGW2-A* and *-D* proteins. Unfortunately, in this study, the full-length cDNAs of *TaGW2-A* and *-B* and only a partial sequence of *TaGW2-D* were obtained. Biochemical analyses therefore focused on the functional characterization of the *TaGW2-A* protein.

TaGW2-A is a functional RING-type E3 ligase

GW2 was previously characterized as a functional E3 RING ligase, *in vitro*. To determine whether *TaGW2-A* is also a functional E3 ligase, an *in vitro* ubiquitination assay was thus performed. A 6×His-*TaGW2-A* fusion protein (52.2 kDa) was produced in *Escherichia coli* and purified using a histidine affinity column. *TaGW2-A* displayed self-ubiquitination in the presence of Ub, E1, and E2 in the reaction mixture, as shown by the high molecular weight bands of ubiquitinated protein detected by western blot analysis using an anti-ubiquitin antibody (Fig. 3, lane 5). The molecular weight of the bands matched the molecular weight of the 6×His-*TaGW2-A* fusion protein, with one Ub moiety (60.2 kDa) or more. The same pattern of bands was observed for rice *GW2* tested under the same reaction conditions (data not shown). *TaGW2-A* autoubiquitination was abolished if the E1 or E2 enzymes were absent (Fig. 3, lanes 1–4). These results thus show that *TaGW2-A* displays E3 ubiquitin ligase activity *in vitro*, as does its rice counterpart.

TaGW2-A is a nucleocytoplasmic protein and is exported from the nucleus via a functional LxxLxL-type nuclear export signal

To gain insight into the cellular function of *TaGW2-A*, its sub-cellular localization was next examined. Confocal microscopy experiments were performed to assess the distribution of *TaGW2-A* fused to yellow fluorescent protein (YFP) (Fig. 4A). YFP-*TaGW2* was transiently co-expressed by agroinfiltration in tobacco leaf epidermal cells, with cytosolic cyan fluorescent protein (cCFP) as a control. Both proteins were observed in the cytosol and the nucleus (Fig. 4B, merged image). YFP-*TaGW2* was also detected in the nucleolus (Fig. 4B, merged image, arrow). The nuclear localization of cCFP results from passive diffusion through nuclear pore complexes because of its small size (22.3 kDa). As passive diffusion is generally limited to proteins less than ~40 kDa (Wagstaff and Jans, 2009), the size of YFP-*TaGW2* (69.6 kDa) excludes its passive nuclear diffusion, indicating that *TaGW2-A* has a genuine nuclear localization.

The fact that *TaGW2-A* was detected in both the cytoplasm and the nucleus prompted a search for *TaGW2-A* motifs that may govern its intracellular localization. It remains unclear how *TaGW2-A* enters the nucleus, as no nuclear localization signal has been yet identified in the primary sequence of *TaGW2-A*. However, a leucine-rich region (³⁴LxxLxL³⁹) was identified that might act as an NES. To determine whether this putative NES was functional, deletions and mutations of *TaGW2-A*

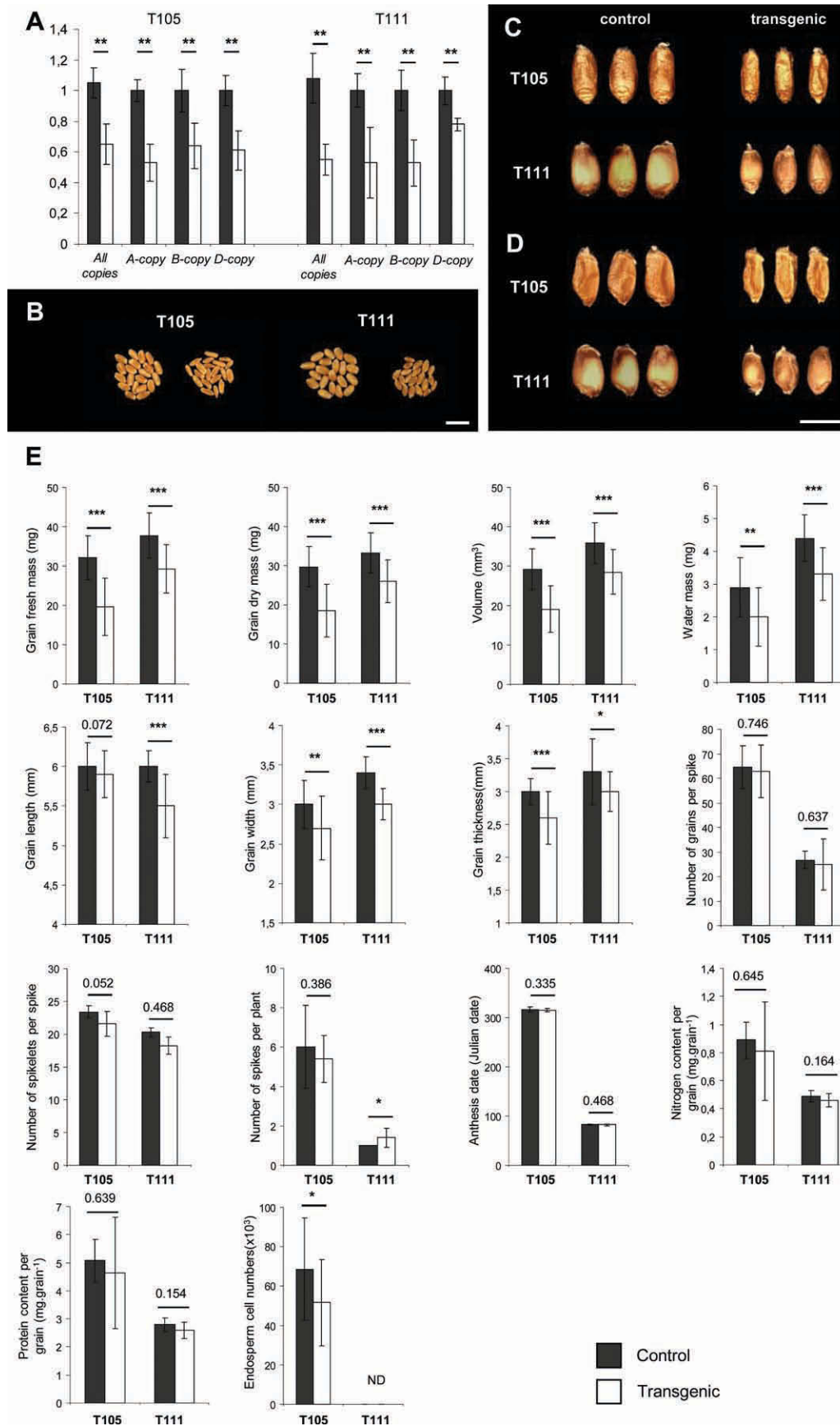


Fig. 1. Molecular characterization and phenotyping for grain yield-related traits of T_2 transgenic and null-segregant control plants from T105 and T111 events. (A) Relative transcript level of endogenous *TaGW2* homoeologues in T_2 transgenic ($n=21$ and $n=8$, in T105 and T111 lines, respectively, grey) and null-segregant control plants ($n=5$ and $n=9$, in T105 and T111 lines, respectively, black). Quantification

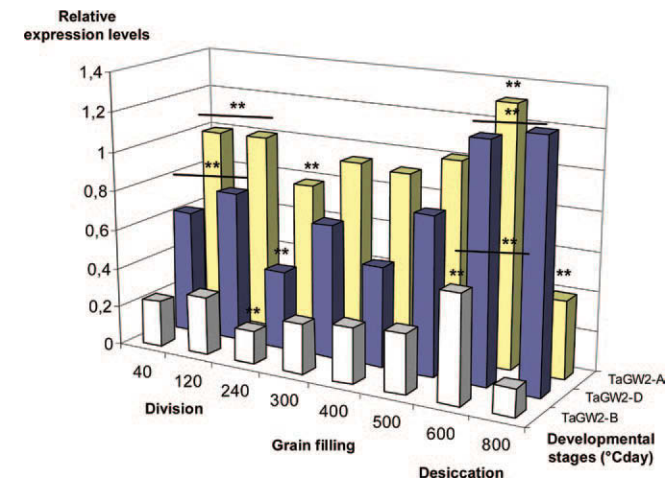


Fig. 2. Relative *TaGW2* homoeologue mRNA levels at different grain development stages. Time is measured in cumulative growth degree days (°Cd). Quantification was performed by qRT-PCR. Relative expression values were calculated using the 2- $\Delta\Delta C_t$ method, using *Ta2776* as an internal control, and the 40 °Cd stage as a calibrator. All data are given as the mean of three independent biological replicates. Comparisons of transcript accumulation between grain developmental stages were performed independently for each copy using Scheffe's multiple comparison test. Asterisks indicate that means are significantly different at $P < 0.01$. (This figure is available in colour at JXB online.)

fused to YFP were generated (Fig. 4A). In the YFP-mNE-*STaGW2* substitution mutant, the three leucines were replaced by alanine (³⁴AxxAxA³⁹). Consistent with a disruption of NES function, YFP-mNE*STaGW2* accumulated exclusively in the nucleus (Fig. 4C). This result indicates that the LxxLxL motif is essential for nuclear export of YFP-*TaGW2*, and that disruption of this motif led to nuclear retention of the chimeric protein. Subsequently, the first 39 amino acids of *TaGW2*-A, which include the putative LxxLxL-type NES, were investigated to be used to export cYFP from the nucleus [(1-39)*TaGW2*-YFP]. The cYFP protein alone, like cCFP, accumulated in the cytosol and diffused passively into the nucleus (Fig. 4D). (1-39) *TaGW2*-YFP was mainly located in the cytosol (Fig. 4E), although the size of this fusion protein (26.83 kDa) would still allow it to diffuse passively into the nucleus. Thus amino acids 1-39 of *TaGW2*-A were sufficient to export cYFP from

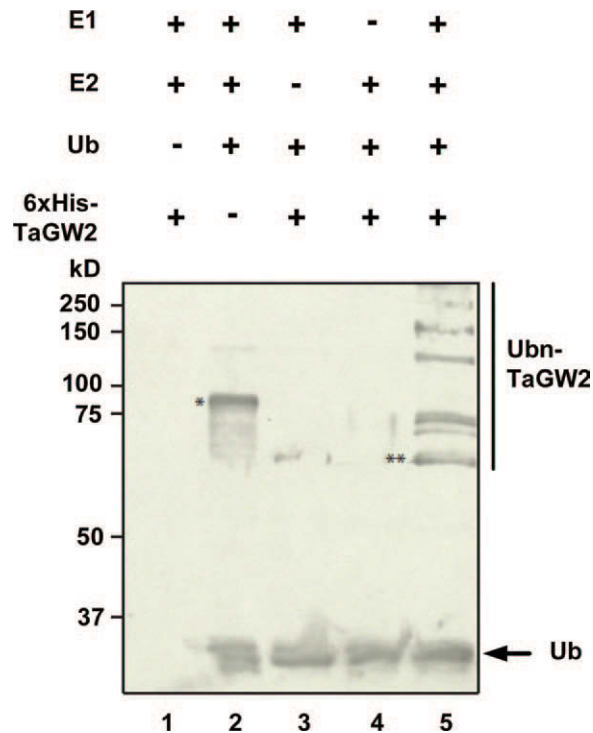


Fig. 3. E3 ubiquitin ligase activity of *TaGW2*-A. 6xHis-*TaGW2*-A was assayed for E3 activity in the presence or absence of human E1 (A1S9), E2 [glutathione S-transferase (GST)-Ubch5b], and Ub (GST-Ub). Ubiquitin and ubiquitinated proteins were detected by western blotting using anti-ubiquitin antibody (P4D1, Santa Cruz). An asterisk indicates monoUb-E2 intermediates; in the absence of E3, ubiquitin cannot be transferred to E3 and remains bound to the E2 enzyme. Double asterisks indicate monoUb-*TaGW2*-A.

the nucleus. However, amino acids 19-29 of *TaGW2*-A [(19-39) *TaGW2*-YFP] were not sufficient to export cYFP from the nucleus (Fig. 4F). This result shows that the LxxLxL motif alone in the N-terminal part of *TaGW2*-A was not sufficient to shuttle cYFP out of the nucleus and further suggests the probable importance of the three-dimensional conformation, which makes the LxxLxL motif accessible.

Interestingly, YFP-mNE*STaGW2* was excluded from the nucleolus, which suggests that the NES might also act as a retention signal in the nucleolus. Indeed, the fusion of *TaGW2*-A amino acids 19-39 to YFP clearly resulted in an accumulation of the fusion protein in the nucleolus (Fig. 4F, indicated by the

was performed by qRT-PCR. Relative expression values were calculated using the 2- $\Delta\Delta C_t$ method, using *Ta54227* as an internal control, and the mean of $\Delta\Delta C_t$ values of the null segregants from the transgenic event as a calibrator. Data are given as the mean \pm SD. Asterisks indicate that means are significantly different at $P < 0.01$ (Scheffe's test). (B-D) Mature grains of transgenic and control plants from T_2 progeny of T105 and T111 lines. (B) Twenty grains from one whole spike of transgenic (right) and control (left) plants from both transgenic lines. Bar=1 cm. Dorsal (C) and lateral (D) views of basal grains from central spikelets of main spikes. Bar=5 mm. (E) Characterization of grain yield-related traits of T_2 transgenic lines. Mean \pm SD. All parameters except endosperm cell numbers were measured on grains harvested at full maturity (1000 °Cd). Endosperm cell number counting was performed on grains harvested at 260 °Cd. *** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$, respectively. The numbers of T_2 transgenic and null-segregant control plants are $n=21$ and $n=5$, respectively, for the T105 event, and $n=8$ and $n=9$, respectively, for the T111 event. Anthesis date is expressed as the Julian date of the year. The number of days between sowing and anthesis were 144 d and 126 d for T105 and T111, respectively.

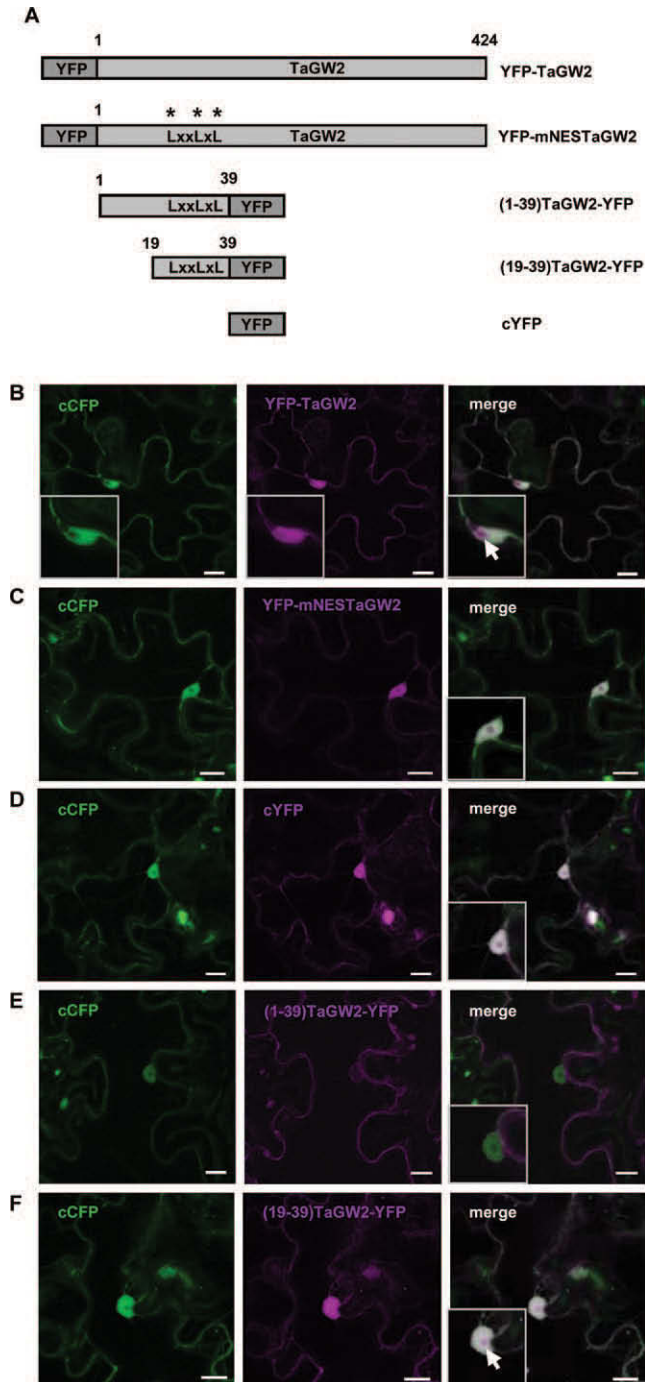


Fig. 4. *TaGW2-A* is a nucleocytoplasmic and nucleolar protein with a functional NES domain. (A) Schematic representation of fusion proteins. (B–F) CLSM images of *Nicotiana tabacum* leaf epidermal cells co-expressing cCFP (left panels) and YFP fusions of full-length or fragments of *TaGW2-A* (middle panels). The merged images are in the right-hand panels. A detailed view of the nucleus is given in the insets. YFP–*TaGW2* (B) co-localizes with cCFP in the cytosol and the nucleus. YFP–*TaGW2* also accumulates in the nucleolus (indicated by an arrow, in the inset). YFP–mNESTaGW2 (C) is retained in the nucleus. cYFP (D) accumulates in both the nucleus and the cytosol. (E) The first 39 amino acids of *TaGW2-A* are able to relocate the cYFP protein from the nucleus to the cytosol. (F) The addition of *TaGW2-A* amino acids 19–39 in the N-terminal part

arrow). However, the first 39 amino acids of *TaGW2-A* fused to YFP did not lead to a nucleolar accumulation of (1–39) *TaGW2*–YFP (Fig. 4E, middle panel). It should be remembered that (1–39)*TaGW2*–YFP was actively exported from the nucleus and so was mainly located in the cytosol. It can be hypothesized that (1–39)*TaGW2*–YFP could not accumulate in the nucleolus because the remaining proportion of (1–39)*TaGW2*–YFP in the nucleoplasm was too small to be addressed to the nucleolus.

Discussion

Silencing of TaGW2 impacts wheat grain size

To summarize, RNAi-mediated down-regulation of *TaGW2* in wheat resulted in significant decreases in final grain fresh, dry and water masses, and grain volume, width, and thickness. No significant difference was observed in the number of grains per spike. A negative correlation between grain number per spike and grain size is frequently observed; the reduced grain dry mass in transgenic plants thus does not result from a larger number of grains per spike (Fisher *et al.*, 1977). Furthermore, no significant differences in the number of spikelets per spike, the number of spikes per plant (in the T105 line), or the anthesis date were observed in transgenic plants compared with the controls, suggesting that the transgene exerted no effect on vegetative traits of the plants and that a direct intrinsic effect of the transgene on grain development was therefore likely.

Two studies previously investigated the genetic association between *TaGW2-A* and wheat yield components (Su *et al.*, 2011; Yang *et al.*, 2012). Although both studies suggest *TaGW2-A* as a negative regulator of grain width, like *GW2* in rice, they diverge as regards the nature of the causal mechanism. By comparing the three homoeologous sequences of *TaGW2* between small and large grain varieties, Su *et al.* (2011) did not find any variation in the coding regions. Instead, a base substitution in the promoter region of *TaGW2-A* was found to be associated with grain width and TGW. Moreover, the authors established a negative correlation between the expression level of *TaGW2* and grain width, in immature grains at 10 d post-flowering. Yang *et al.* (2012) did not report any variation for the coding sequences from chromosomes 6B and 6D but identified a single T base insertion in the eighth exon of *TaGW2-A* of a large grain variety, Lankadali. This insertion leads to a non-functioning protein, which is similar to the deletion in rice *GW2*. This marker was found associated with grain width and TGW in an F₂ population derived from a cross between Lankadali and Chinese Spring. The present study shows that overall down-regulation of *TaGW2* copies results in smaller wheat kernels, thus suggesting that *TaGW2* is a positive regulator of grain size-related traits. The results are notably consistent with those of Su *et al.* (2011) as both imply a modulation of the expression level. However, there are discrepancies between the present results and the results reported by

of YFP results in an accumulation of (19–39)*TaGW2*–YFP in the nucleolus. Its accumulation in the cytosol and the nucleoplasm remains unchanged in comparison with cYFP. Magnification ×20. Bar=10 μm.

Su et al. (2011) and Yang et al. (2012). Three possible hypotheses may explain these discrepancies. One is that the previous studies focused on the A copy of *TaGW2* and, to date, nothing is known as regards the other two copies of the gene. Here, RNAi enabled silencing of the three copies of *TaGW2* at a time. Thus a potential cumulative effect of silencing the three *TaGW2* homoeologues during the whole developmental cycle of the plant may have been observed. A second explanation is that the previous studies dealt with genetic associations between polymorphic markers within the A copy of *TaGW2* and yield components. However, both studies indicated that *TaGW2-A* is located in a genomic region with a major yield QTL, close to the 6A centromere (Snape et al., 2007). Therefore, the observed genetic association could be explained by linkage disequilibrium between these markers and a gene not yet identified with a large effect on yield components. Interestingly, Li et al. (2010a) conducted similar studies in maize and identified the favourable alleles (i.e. the relative contribution of each parental allele) using either an association analysis or a QTL mapping approach. The two approaches provided conflicting results and the authors suggested that the polymorphic marker they used could be in linkage disequilibrium with the actual functional polymorphism (Li et al., 2010a). A third possibility is off-target effects caused by the long RNAi hairpin used in this study to down-regulate *TaGW2* genes. Although a search in URGI and CerealsDB wheat genomic databases did not reveal any 20 nucleotide identity stretches in coding sequences with known function, it cannot be excluded that the RNAi hairpin silences additional genes which positively impact grain size-related traits.

Although similar results were obtained in both transgenic lines, interestingly the effect of *TaGW2* down-regulation on grain yield-related traits was always less substantial in T111 transgenic plants than in T105. In both transgenic lines, the A and B copies of *TaGW2* were silenced to similar levels. However, *TaGW2-D* transcript accumulation was reduced by almost twice as much in the T105 line as in T111. This difference in *TaGW2-D* silencing could thus explain the different magnitudes of the observed effects in both transgenic lines. Moreover, this result suggests

that the different homeologous genes contribute to this quantitative trait, and that any modulation of the expression level of one copy could quantitatively impact the yield components.

Various approaches have been attempted to determine grain weight potential in wheat and other crops. Final grain dry weight, which results from an accumulation of dry matter, is mainly governed by the number of endosperm cells and their expansion (Hoshikawa, 1962; Brocklehurst, 1977; Vilhar et al., 2002). Water uptake is a major determinant of grain size during the early phase of wheat grain development (Schnyder and Baum, 1992; Nadaud et al., 2010). Indeed, the initial growth of grains is mostly due to a rapid net deposition of water in the grain (Schnyder and Baum, 1992). Schnyder and Baum (1992) proposed that water content largely determines volumetric growth of the grain during endosperm cell division. Other studies support the hypothesis that grain weight determination is driven by early growth of the pericarp, which restricts grain expansion and consequently grain volume (Calderini et al., 1999a, b; Calderini and Reynolds, 2000; Garcia et al., 2005; Ugarte et al., 2007). During the grain-filling stage, grain dry weight increases, with starch accumulating between 13 and 34 DPA and storage proteins between 10 and 36 DPA (Schnyder and Baum, 1992; Dupont and Altenbach, 2003). Moreover, during this stage, water and dry matter dynamics are closely related (Schnyder and Baum, 1992; Calderini and Reynolds, 2000), and the end of water accumulation in the endosperm coincides with the end of dry matter accumulation (Schnyder and Baum, 1992). Finally, during grain desiccation, the endosperm tissue undergoes programmed cell death and the kernel rapidly loses all but 10–15% of its water content.

The expression patterns of *TaGW2-A* and *-D* described in this study were consistent with a role for *TaGW2* in regulating endosperm cell number. Based on the RNAi results and on current knowledge of the determination of grain weight potential in wheat, a model for the role of *TaGW2* in the regulation of grain size and grain weight in wheat can be proposed (Fig. 5). For the time being, it is difficult to identify clearly which homoeologous copy acts in these processes, so the overall *TaGW2* will

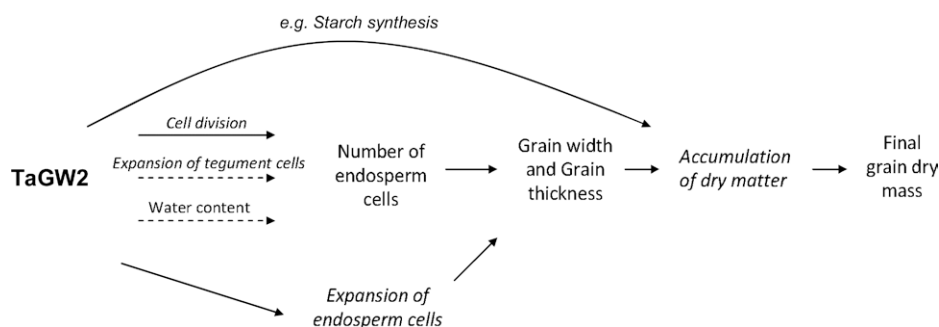


Fig. 5. Proposed model for the role of *TaGW2* in the regulation of grain size and grain weight in wheat. *TaGW2* could positively regulate the number and expansion of endosperm cells, thus influencing both grain width and grain thickness, leading to enhanced accumulation of dry matter and ultimately increasing final grain dry mass. The action of *TaGW2* on endosperm cell numbers might be direct by controlling cell division in the endosperm or indirect by establishing volumetric restrictions on endosperm growth via control of the expansion of tegument cells or of water uptake into the grain. *TaGW2* direct control of the accumulation of dry matter cannot be excluded. Processes are indicated in italics. Solid and dashed arrows indicate a presumed *TaGW2* direct or indirect effect on the trait, respectively.

be considered in this model. *TaGW2* may positively regulate the number and expansion of endosperm cells, thus influencing both grain width and grain thickness, leading to enhanced accumulation of dry matter and ultimately increasing final grain dry mass. *TaGW2* may control endosperm cell number directly by controlling endosperm cell division or indirectly by establishing volumetric restrictions on endosperm growth by controlling the expansion of tegument cells or water uptake into the grain. It can also not be excluded that *TaGW2* directly regulates the accumulation of dry matter during the grain-filling period through the control of starch synthesis.

TaGW2 in wheat and GW2 in rice have divergent functions in grain development

Interestingly, the grain traits (grain weight, width, and thickness) affected by the RNAi of *TaGW2* were the same as those altered in the *GW2* antisense lines and in the WY3 rice variety, which encodes a loss-of-function variant of *GW2* (Song *et al.*, 2007). However, *TaGW2* acts positively on these traits, opposite to rice *GW2*, which acts negatively. In addition, no significant differences in grain number per spike and in the number of spikes per plant in wheat were observed, but *GW2* loss-of-function in rice reduced the grain number per spike by 29.9% and increased the panicle number per plant by ~27%. It is also noteworthy that in rice, *GW2* regulates the number of cells in the spikelet hull but has no effect on the number of cells in the endosperm; in contrast, *TaGW2* regulates endosperm cell number, at least in the T105 line. These two homologous genes thus might function differently during grain development in wheat and rice. This discrepancy of function between two orthologous genes was unexpected as orthologous genes often have conserved functions (Fan *et al.*, 2006; Li *et al.*, 2010b). Interestingly, Distelfeld *et al.* (2012) reported that, despite their orthologous relationship, the wheat *GPC* genes and the rice *Os07g37920* gene have divergent functions. The wheat *GPC* genes control senescence, but the rice orthologous *Os07g37920* gene controls anther dehiscence. Furthermore, divergence of function between several orthologues has also been reported between human and mouse (Gharib and Robinson-Rechavi, 2011).

In this study, biochemical and functional characterizations of the TaGW2-A protein demonstrated first that TaGW2-A is a functional E3 ligase, like its rice counterpart, and secondly that TaGW2-A has genuine nucleolar, nuclear, and cytoplasmic localizations. A LxxLxL-type NES, at the N-terminus, was also characterized as necessary to shuttle TaGW2-A out of the nucleus, and the region between amino acids 1 and 39 of TaGW2-A was shown to be sufficient for this export. It was also observed that this NES LxxLxL motif acts as a signal to retain TaGW2-A in the nucleolus. The RING domain and the LxxLxL motif in the amino acid sequence of TaGW2-B and the TaGW2-D sequence from Chinese Spring (Supplementary Fig. S1 at JXB online) are strictly conserved and it could thus be expected that TaGW2-B and -D are also functional E3s and are nucleocytoplasmic, although this remains to be validated.

Two hypotheses might be proposed as regards the divergence of functions between rice *GW2* and *TaGW2*. One hypothesis is that they may interact with different targets. GW2 and TaGW2-A

share 87% overall amino acid sequence identity. They are highly conserved in their N-terminal regions (93.8% sequence identity from amino acid 1 to 180) but diverge more in their C-terminal regions (82.5% sequence identity from amino acid 181 to 424), where Song *et al.* (2007) suggested that a substrate-binding domain is located (Supplementary Fig. S2 at JXB online). These amino acid substitutions may alter key residues for interaction with target substrate or other protein partners of the signalling pathway. A second hypothesis is that the different subcellular partitioning of GW2 and its wheat orthologue could explain the divergence in the observed phenotypes. Indeed, Song *et al.* (2007) showed that GW2 displays a cytosolic localization whereas TaGW2-A accumulates in the cytosol, the nucleoplasm, and the nucleolus. Gharib and Robinson-Rechavi (2011) notably referred to divergent subcellular localization as a source of variations in orthologue function between human and mouse. Further examination of TaGW2 function, including identification of its interacting proteins, will both shed light on the differences compared with rice GW2 and provide useful insights on the mechanisms of regulation of grain size and thus yield in important crop species.

GenBank accession numbers

The GenBank accession numbers of *TaGW2-A* and *-B* are JN896622 and JN896623, respectively.

Supplementary data

Supplementary data are available at JXB online.

Supplementary information SI. Detailed Materials and methods.

Figure S1. Nucleotide and putative amino acid alignments of the TaGW2-A, -B, and -D CDS.

Figure S2. Amino acid alignment of GW2 and TaGW2-A.

Table S1. List of the primers used in this study.

Acknowledgments

This work was supported by a grant from the French Ministry for Higher Education and Research. The authors acknowledge Nicole Allard and Joëlle Messaoud for help in phenotyping the transgenic experiments; Krzysztof Jagla and Caroline Vachias for advice on the microscopy platform in Joint Research Unit UMR6247 GRd; Mireille Dardevet and Monique Sibaud for technical assistance; Dominique Marcon for images of transgenic grains; and Nelly Lajoinie for growing the plants.

References

- Andersson A, Johansson E, Oscarson P. 2004. Post-anthesis nitrogen accumulation and distribution among grains in spring wheat spikes. *Journal of Agricultural Science* **142**, 525–533.
- Boulafous A, Saint-Jore-Dupas C, Herranz-Cardo MC, Pagny-Salehabadi S, Plasson C, Garidou F, Kiefer-Meyer MC, Ritzenthaler C, Faye L, Gomord V. 2009. Cytosolic N-terminal

arginine-based signals together with a luminal signal target a type II membrane protein to the plant ER. *BMC Plant Biology* **9**, 144.

Brocklehurst PA. 1977. Factors controlling grain weight in wheat. *Nature* **266**, 348–349.

Calderini DF, Abeledo LG, Savin R, Slafer GA. 1999a. Effect of temperature and carpel size during pre-anthesis on potential grain in wheat. *Journal of Agricultural Science, Cambridge* **132**, 453–459.

Calderini DF, Abeledo LG, Savin R, Slafer GA. 1999b. Final grain weight in wheat as affected by short periods of high temperature during pre- and post-anthesis under field conditions. *Australian Journal of Plant Physiology* **26**, 453–458.

Calderini DF, Ortiz-Monasterio I. 2003. Grain position affects grain macronutrient and micronutrient concentrations in wheat. *Crop Science* **43**, 141–151.

Calderini DF, Reynolds MP. 2000. Changes in grain weight as a consequence of de-graining treatments at pre- and post-anthesis in synthetic hexaploid lines of wheat. *Australian Journal of Plant Physiology* **27**, 183–191.

Capron D, Mouzeyar S, Boulaflois A, Girousse C, Rustenholz C, Laugier C, Paux E, Bouzidi MF. 2012. Transcriptional profile analysis of E3 ligase and hormone-related genes expressed during wheat grain development. *BMC Plant Biology* **12**, 35.

Distelfeld A, Pearce SP, Avni R, Scherer B, Uauy C, Piston F, Slade A, Zhao R, Dubcovsky J. 2012. Divergent functions of orthologous NAC transcription factors in wheat and rice. *Plant Molecular Biology* **78**, 515–524.

Dupont FM, Altenbach SB. 2003. Molecular and biochemical impacts of environmental factors on wheat grain development and protein synthesis. *Journal of Cereal Science* **38**, 133–146.

Fan CC, Xing YZ, Mao HL, Lu TT, Han B, Xu CG, Li XH, Zhang QF. 2006. GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theoretical and Applied Genetics* **112**, 1164–1171.

Fisher RA, Aguilar I, Laing DR. 1977. Post-anthesis sink size in a high-yielding dwarf wheat: yield response to grain number. *Australian Journal of Agricultural Research* **28**, 165–175.

Garcia D, Saingery V, Chambrier P, Mayer U, Jürgens G, Berger F. 2003. Arabidopsis haiku mutants reveal new controls of seed size by endosperm. *Plant Physiology* **131**, 1661–1670.

Gharib WH, Robinson-Rechavi M. 2011. When orthologs diverge between human and mouse. *Briefings in Bioinformatics* **12**, 436–441.

Hoshikawa K. 1962. Studies on the ripening of wheat grain. 4. Influence of temperature upon the development of the endosperm. *Proceedings of the Crop Science Society of Japan* **30**, 228–231 (in Japanese with English summary).

Kato K, Miura H, Sawada S. 2000. Mapping QTLs controlling grain yield and its components on chromosome 5A of wheat. *Theoretical and Applied Genetics* **101**, 1114–1121.

Li Q, Li L, Yang X, Warburton ML, Bai G, Dai J, Li J, Yan J. 2010a. Relationship, evolutionary fate and function of two maize co-orthologs of rice GW2 associated with kernel size and weight. *BMC Plant Biology* **10**, 143.

Li Q, Yang XH, Bai GH, Warburton ML, Mahuku G, Gore M, Dai JR, Li JS, Yan JB. 2010b. Cloning and characterization of a putative GS3 ortholog involved in maize kernel development. *Theoretical and Applied Genetics* **120**, 753–763.

McCartney CA, Somers DJ, Humphreys DG, Lukow O, Ames N, Noll J, Cloutier S, Mc Callum BD. 2005. Mapping quantitative trait loci controlling agronomic traits in the spring wheat cross RL4452×AC domain. *Genome* **48**, 870–883.

Nadaud I, Girousse C, Debiton C, Chambon C, Bouzidi MF, Martre P, Branlard G. 2010. Proteomic and morphological analysis of early stages of wheat grain development. *Proteomics* **10**, 2901–2910.

Paolacci AR, Tanzarella OA, Porceddu E, Ciaffi M. 2009. Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. *BMC Molecular Biology* **10**, 11.

Pellegrineschi A, Noguera LM, Skovmand B, Brito RM, Velazquez L, Salgado MM, Hernandez R, Warburton M, Hoisington D. 2002. Identification of highly transformable wheat genotypes for mass production of fertile transgenic plants. *Genome* **45**, 421–430.

Pfaffl MW. 2004. Quantification strategies in real-time PCR. In: Bustin SA, ed. *The real-time PCR encyclopedia. A–Z of quantitative PCR*. International University Line, 87–120.

Röder MS, Huang XQ, Börner A. 2008. Fine mapping of the region on wheat chromosome 7D controlling grain weight. *Functional and Integrative Genomics* **8**, 79–86.

SAS Institute. 2002. SAS/GRAPH Software: Reference Volume 1, Version 8. Cary, NC: SAS Institute, Inc.

Schmittgen TD, Livak KJ. 2008. Analyzing real-time PCR data by the comparative C(T) method. *Nature Protocols* **3**, 1101–1108.

Schnyder H, Baum U. 1992. Growth of the grain of wheat (*Triticum aestivum* L.). The relationship between water content and dry matter accumulation. *European Journal of Agronomy* **1**, 51–57.

Sears ER. 1966. Nullisomic-tetrasomic combinations in hexaploid wheat. In: Riley R, Lewis KR, eds. *Chromosome manipulations and plant genetics*. Edinburgh: Oliver and Boyd, 29–45.

Simmonds DH, O'Brien TP. 1981. Morphological and biochemical development of wheat endosperm. In: Pomerantz Y, ed. *Advances in cereal science and technology*. St Paul, MN: American Association of Cereal Chemists, 5–70.

Singh BK, Jenner CF. 1982. Association between concentrations of organic nutrients in the grain, endosperm cell number and grain dry weight within the ear of wheat. *Australian Journal of Plant Physiology* **9**, 83–95.

Snape JW, Foulkes MJ, Simmonds J, Leverington M, Fish LJ, Wang YK, Ciavarella M. 2007. Dissecting genexenvironmental effects on wheat yields via QTL and physiological analysis. *Euphytica* **154**, 401–408.

Song XJ, Huang W, Shi M, Zhu MZ, Lin HX. 2007. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nature Genetics* **39**, 623–630.

Sosulski FW, Imafidon GI. 1990. Amino acid composition and nitrogen-to-protein conversion factors for animal and plant foods. *Journal of Agricultural and Food Chemistry* **38**, 1351–1356.

- Su Z, Hao C, Wang L, Dong Y, Zhang X.** 2011. Identification and development of a functional marker of TaGW2 associated with grain weight in bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* **122**, 211–223.
- Ugarte C, Calderini DF, Slafer GA.** 2007. Grain weight and grain number responsiveness to pre anthesis temperature in wheat, barley and triticale. *Field Crops Research* **100**, 240–248.
- Viestra RD.** 2009. The ubiquitin–26s proteasome system at the nexus of plant biology. *Nature Reviews Molecular and Cellular Biology* **10**, 385–397.
- Vilhar B, Kladnik A, Blejec A, Chourey PS, Dermastia M.** 2002. Cytometrical evidence that the loss of seed weight in the miniature1 seed mutant of maize is associated with reduced mitotic activity in the developing endosperm. *Plant Physiology* **129**, 23–30.
- Wagstaff KM, Jans DA.** 2009. Importins and beyond: non-conventional nuclear transport mechanisms. *Traffic* **10**, 1188–1198.
- Weng JF, Gu SH, Wan XY, et al.** 2008. Isolation and initial characterization of GW5, a major QTL associated with rice grain width and weight. *Cell Research* **18**, 1199–1209.
- Wright M, Dawson J, Dunder E, Suttie J, Reed J, Kramer C, Chang Y, Novitzky R, Wang H, Artim-Moore L.** 2001. Efficient biolistic transformation of maize (*Zea mays* L.) and wheat (*Triticum aestivum* L.) using the phosphomannose isomerase gene, *pmi*, as the selectable marker. *Plant Cell Reports* **20**, 429–436..
- Yang Z, Bai Z, Li X, Wang P, Wu Q, Yang L, Li L, Li X.** 2012. SNP identification and allelic-specific PCR markers development for TaGW2, a gene linked to wheat kernel weight. *Theoretical and Applied Genetics* **125**, 1057–1068.

ANNEXES

Annexe 1

COMPOSITION DES MILIEUX ET SOLUTIONS

Milieu BY-2 liquide

	<i>1L</i>
Murashige et Skoog	4,3g
Saccharose	30g
Thiamine (solution stock à 1 g/L)	1 ml
Myoinositol (solution stock à 51 g/L)	1 ml
2,4-D (solution stock à 0,2 g/L)	1 ml
KH ₂ PO ₄ (solution stock à 20 g/L)	5 ml
H ₂ O	Qsp 1L

pH 5,2 (pour la culture de cellules en suspension)
ou pH 5,8 (pour la culture de cals en milieu gélosé)
Autoclave

Milieu BY-2 solide

Milieu BY-2 liquide + 1,2% d'agar
Autoclave

Milieu Luria-Bertani (LB)

	<i>1L</i>
Bacto-tryptone	10g
Bacto-yeast extract	5g
NaCl	10g
H ₂ O	Qsp 1L

Autoclave

MILIEUX DE LEVURES

0,5X YPDA

	<i>100 ml</i>	<i>200 ml</i>	<i>1L</i>
YPD	2,5g	5g	25g
Adenine hémisulfate	4 mg	8 mg	40 mg
H ₂ O	Qsp 100 ml	Qsp 200 ml	Qsp 1L

pH 6
Autoclave

1X YPDA

	<i>100 ml</i>
YPD	5g
Adenine hémisulfate	8 mg
H ₂ O	Qsp 100 ml

pH 6
Autoclave

2X YPDA

YPD

Adenine hémisulfate

H₂O

pH 6

Autoclave

100 ml

10g

16 mg

Qsp 100 ml

200 ml

20g

32 mg

Qsp 200
ml*1L*

100g

160 mg

Qsp 1L

SD/-L

YNB sans sulfate d'ammonium

Ammonium sulfate

Drop out without Leucine (Sigma)

H₂O

pH 5,9

Autoclave

+ 10 ml de Glucose 50% par 250 ml de milieu

200 ml

0,34g

1g

0,32g

Qsp 200 ml

300 ml

0,51g

1,5g

0,48g

Qsp 300
ml*600 ml*

1,02g

3g

0,96g

Qsp 600
ml**SD/-T**

YNB sans sulfate d'ammonium

Ammonium sulfate

Drop out without Tryptophane (Sigma)

H₂O

pH 5,9

Autoclave

+ 10 ml de Glucose 50% par 250 ml de milieu

100 ml

0,17g

0,5g

0,19g

Qsp 100 ml

SD/-L/-T (ou DDO)

YNB sans sulfate d'ammonium

Ammonium sulfate

Drop out without Leucine and Tryptophane (Sigma)

H₂O

pH 5,9

Autoclave

+ 10 ml de Glucose 50% par 250 ml de milieu

250 ml

0,43g

1,25g

0,4g

Qsp 250 ml

400 ml

0,68g

2g

0,64g

Qsp 400
ml*2L*

3,4g

10g

3,2g

Qsp 2L

SD/-L/-T/-H

YNB sans sulfate d'ammonium

Ammonium sulfate

Drop out without Leucine, Tryptophane, Histidine and
Adenine (Clontech)

Adenine hemisulfate (solution stock 100X)

H₂O*500 ml*

0,85g

2,5g

0,3g

5 ml

Qsp 500 ml

1L

1,7g

5g

0,6g

10 ml

Qsp 1L

2L

3,4g

10g

1,2g

20 ml

Qsp 2L

pH 5,9
Autoclave
+ 10 ml de Glucose 50% par 250 ml de milieu

Adenine hémisulfate (solution 100X): *10 ml*
Adenine hémisulfate (Sigma) 20 mg

Filtrer, stocker à 4°C

<u>SD/-L/-T/-H/-Ad (ou QDO)</u>	<i>100 ml</i>	<i>400 ml</i>	<i>3L</i>
YNB sans sulfate d'ammonium	0,17g	0,68g	5,1g
Ammonium sulfate	0,5g	2g	15g
Drop out without Leucine, Tryptophane, Histidine and Adenine (Clontech)	0,06g	0,24g	1,8g
H ₂ O	Qsp 100 ml	Qsp 400 ml	Qsp 3L

pH 5,9
Autoclave
+ 10 ml de Glucose 50% par 250 ml de milieu

Pour DDO et QDO, quand /X :
X-alpha-Gal (Clontech, à -20°C) 500 µl par 250 ml de milieu

Pour DDO et QDO, quand /A :
Aureobasidine A (Clontech, à 4°C) 62.5 µl par 250 ml de milieu

Un exemple, pour 250 ml de DDO/X/A :
250 ml de milieu SD/-L/-T (ou DDO)
500 µl de X-alpha-Gal
62,5 µl d'Aureobasidine A

Pour DDO et QDO, quand /+ x mM de 3AT (3-amino triazol):

	<i>Par 250 ml de milieu</i>
1 mM 3AT	21 mg
2,5 mM 3AT	50 mg
5 mM 3AT	105 mg
7,5 mM 3AT	155 mg
10 mM 3AT	210 mg

Pour tous les milieux gélifiés de levures et des bactéries: *250 ml*
Agar 5g

MILIEUX DE TRANSFORMATION DE BLE

Milieu de callogénèse

Murashige et Skoog + micro et macroéléments + vitamines	1L 4,4g
Saccharose	60g
Asparagine	0,15g
Thiamine HCl	0,04g
Agar	8g
2,4-D (solution stock à 1mg/ml)	2,5 ml

pH 5,7 avec du KOH

Milieu plasmolysant

Murashige et Skoog + micro et macroéléments + vitamines	1L 4,4g
Maltose	100g
Asparagine	0,15g
Thiamine HCl	0,04g
Agar	8g
2,4-D (solution stock à 1 mg/ml)	2,5 ml

pH 5,7 avec du KOH

Milieu de régénération sélectif (pour PMI) n°1

Murashige et Skoog + micro et macroéléments + vitamines	1L 4,4g
Saccharose	5g
Mannose	1%
Asparagine	0,15g
Thiamine HCl	0,04g
IAA (solution stock à 1 mg/ml)	0,5 ml
BAP (solution stock à 1 mg/ml)	1 ml
Agar	8g

pH 5,7 avec du KOH

Milieu de régénération sélectif (pour PMI) n°2

Murashige et Skoog + micro et macroéléments + vitamines	1L 4,4g
Saccharose	20g
Mannose	0,5%
Asparagine	0,15g
Thiamine HCl	0,04g
IAA (solution stock à 1 mg/ml)	0,5 ml

BAP (solution stock à 1 mg/ml)	1 ml
Agar	8g

pH 5,7 avec du KOH

<u>Milieu de régénération sélectif (pour PMI) n°3</u>	<i>1L</i>
Murashige et Skoog + micro et macroéléments + vitamines	4,4g
Mannose	1,5%
Asparagine	0,15g
Thiamine HCl	0,04g
IAA (solution stock à 1 mg/ml)	0,5 ml
BAP (solution stock à 1 mg/ml)	1 ml
Agar	8g

pH 5,7 avec du KOH

<u>Milieu d'enracinement</u>	<i>1L</i>
Murashige et Skoog + micro et macroéléments + vitamines	2,2g
Saccharose	20g
IAA (solution stock à 1 mg/ml)	1 ml
Agar	8g

pH 5,7 avec du KOH

<u>PEG/LiOAc</u>	Concentration finale
PEG 3350	40% (p/v)
LiOAc	120 mM

Annexe 2 : Protocole de mise en compétence et de transformation d'*E.coli* souche DH5 α

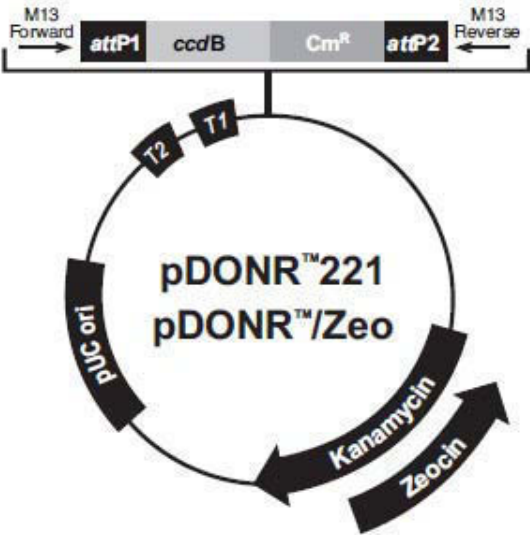
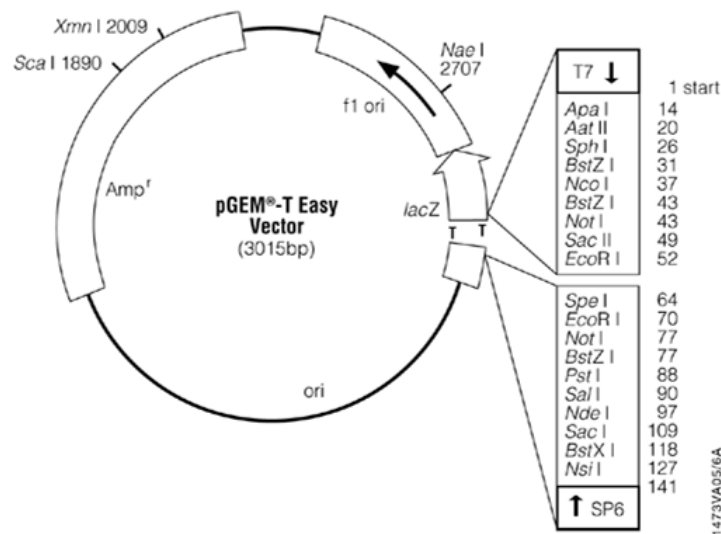
Production de bactéries compétentes

Une colonie isolée de la souche DH5 α d'*E.coli* est inoculée dans 3 ml de milieu LB, et incubée une nuit, à 37°C, sous agitation à 180 rpm. La préculture est alors diluée au 1/1000^{ème} avec du milieu LB frais (soit 150 μ l de préculture dans 150 ml de milieu LB), et incubée de nouveau jusqu'à obtention d'une densité optique à 600 nm de la culture bactérienne de 0,5-0,7. La totalité de la culture est alors répartie dans des tubes stériles de 50 ml, sur la glace. Après centrifugation à 2000g, à 4°C, pendant 7 min, le culot bactérien est repris dans 25 ml de CaCl₂ 0,05 M préalablement refroidi, puis le mélange est incubé pendant 20 min à 4°C et enfin centrifugé pendant 5 min à 1000 g. Le culot est repris dans 4 ml de CaCl₂ 0,05M, et le mélange est incubé une nuit à 4°C. Les bactéries compétentes peuvent être directement utilisées pour leur transformation, ou être stockées à -80°C en présence de glycérol 15% (v/v).

Transformation de bactéries *E.coli* souche DH5 α compétentes

Les bactéries compétentes congelées sont décongelées lentement dans la glace, puis 5 μ l de produit de ligation sont ajoutés à 100 μ l de la suspension bactérienne compétente. Le mélange est incubé pendant 20 min sur la glace, puis soumis au choc thermique suivant : 2 min, à 42°C (bain-marie), 2 min sur la glace. Ensuite, 500 μ l de milieu LB sont ajoutés dans le tube, et la suspension bactérienne transformée est incubée 45 min, à 37°C, sous agitation à 200 rpm. Les bactéries transformées sont centrifugées pendant 3 min à 5000 g, et le culot est finalement repris dans 150 μ l de milieu LB frais. Les bactéries sont étalées sur des boîtes de milieu LB supplémenté en antibiotique ; et les boîtes sont placées une nuit à 37°C.

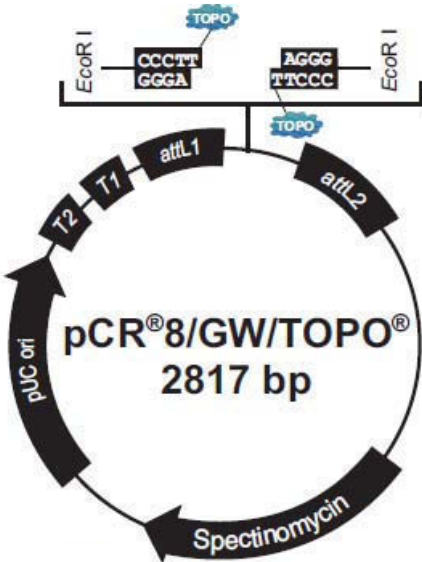
Annexe 3 : Cartes des vecteurs plasmidiques



Comments for:

pDONR™221
4762 nucleotides

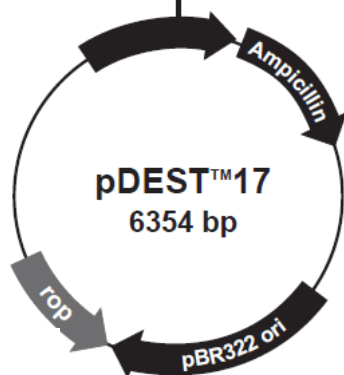
rrnB T2 transcription termination sequence (c):	268-295
rrnB T1 transcription termination sequence (c):	427-470
M13 Forward (-20) priming site:	537-552
attP1:	570-801
ccdB gene (c):	1197-1502
Chloramphenicol resistance gene (c):	1847-2506
attP2 (c):	2754-2985
T7 Promoter/priming site (c):	3000-3019
M13 Reverse priming site:	3027-3043
Kanamycin resistance gene:	3156-3965
EM7 promoter (c):	---
Zeocin resistance gene (c):	---
pUC origin:	4086-4759



Comments for pCR®8/GW/TOPO®
2817 nucleotides

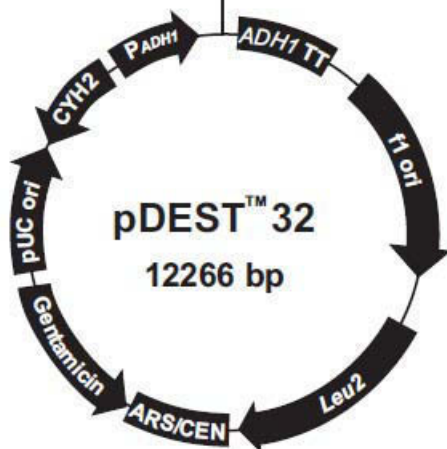
rrnB T2 transcription termination sequence:	bases 268-295
rrnB T1 transcription termination sequence:	bases 427-470
M13 forward (-20) priming site:	bases 537-552
attL1:	bases 569-668
GW1 priming site:	bases 607-631
TOPO® recognition site 1:	bases 678-682
TOPO® recognition site 2:	bases 683-687
attL2:	bases 696-795
GW2 priming site:	bases 733-757
T7 Promoter/priming site:	812-831 (c)
M13 reverse priming site:	bases 836-852
Spectinomycin promoter:	bases 930-1063
Spectinomycin resistance gene (Spn ^R):	1064-2074
pUC origin:	bases 2141-2814

(c) = complementary sequence



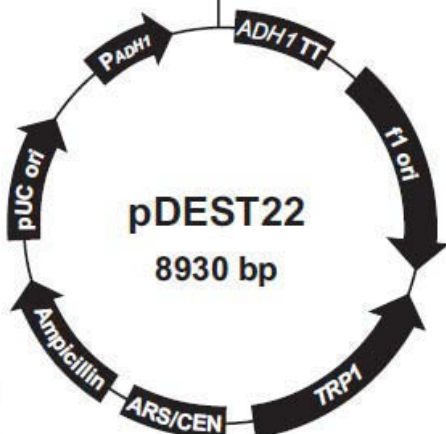
Comments for pDEST™17
6354 nucleotides

T7 promoter: bases 21-40
Ribosome binding site (RBS): bases 86-92
Initiation ATG: bases 101-103
6xHis tag: bases 113-130
attR1: bases 140-264
Chloramphenicol resistance gene (Cm^R): bases 373-1032
ccdB gene: bases 1374-1679
attR2: bases 1720-1844
T7 transcription termination region: bases 1855-1983
bla promoter: bases 2471-2569
Ampicillin (bla) resistance gene: bases 2570-3430
pBR322 origin: bases 3575-4248
ROP ORF: bases 4619-4810 (C)
C=complementary strand



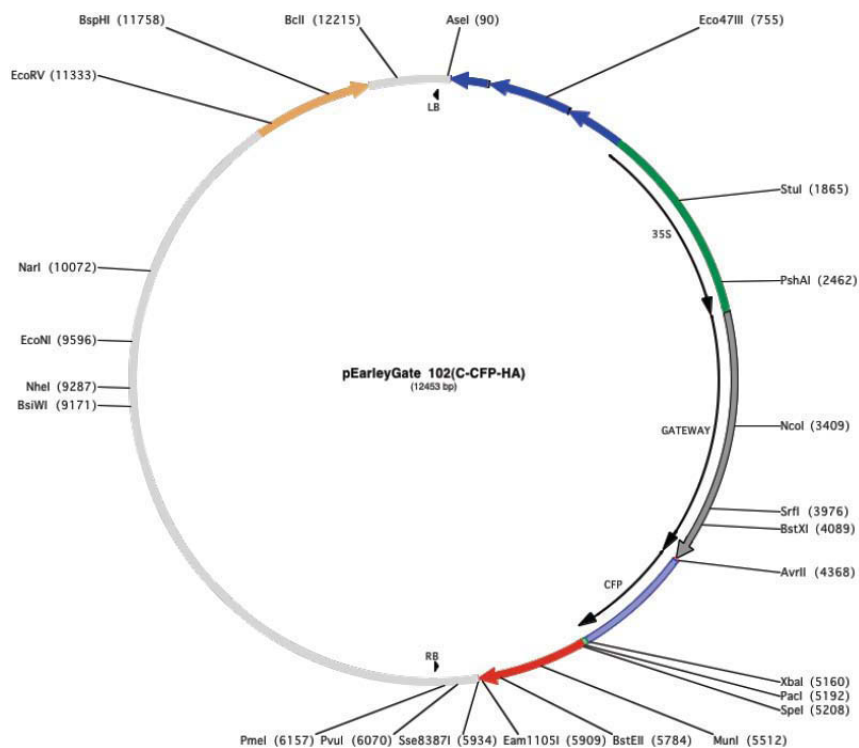
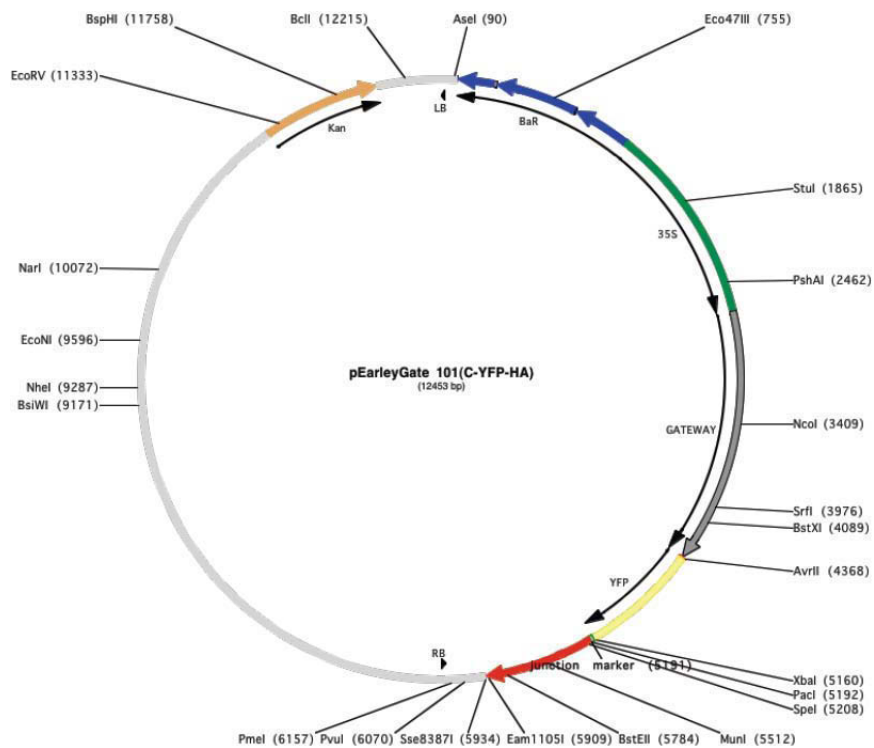
Comments for pDEST™32
12266 nucleotides

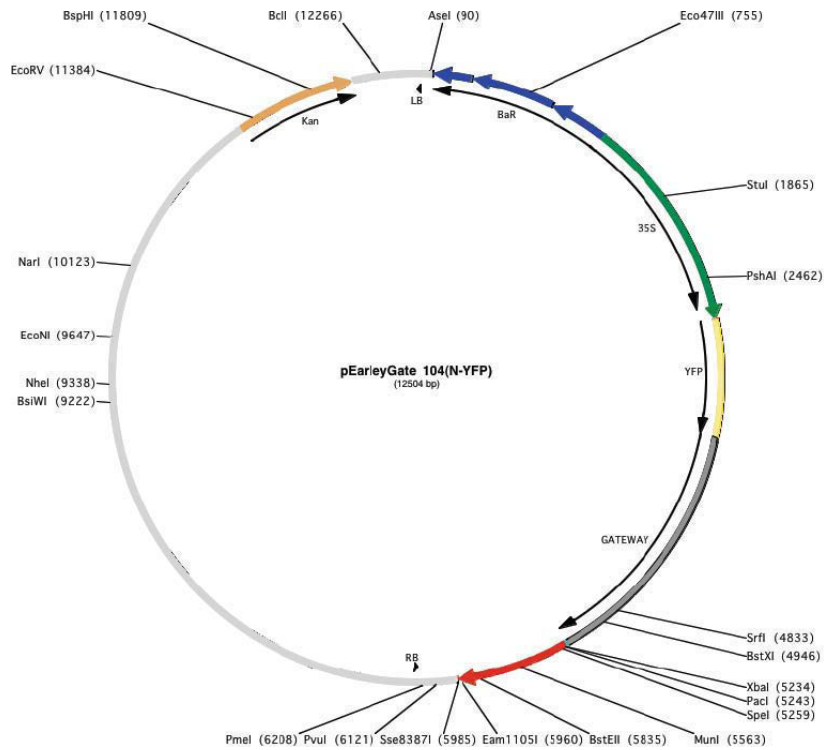
ADH1 promoter: bases 103-1557
GAL4 DNA binding domain: bases 1581-2024
attR1 site: bases 2037-2161
Chloramphenicol resistance (Cm^R) gene: bases 2411-3070
ccdB gene: bases 3411-3716
attR2 site: bases 3757-3881
ADH1 transcription termination region: bases 4119-4276
f1 origin: bases 4603-5058
Leu2 gene: bases 5767-6861
ARS4/CEN6 origin: bases 7589-8107
Gentamicin resistance gene: bases 8452-8985 (c)
pUC origin: bases 9833-10506
Cycloheximide sensitivity (CYH2): bases 11445-11894 (c)
(c) = complementary strand



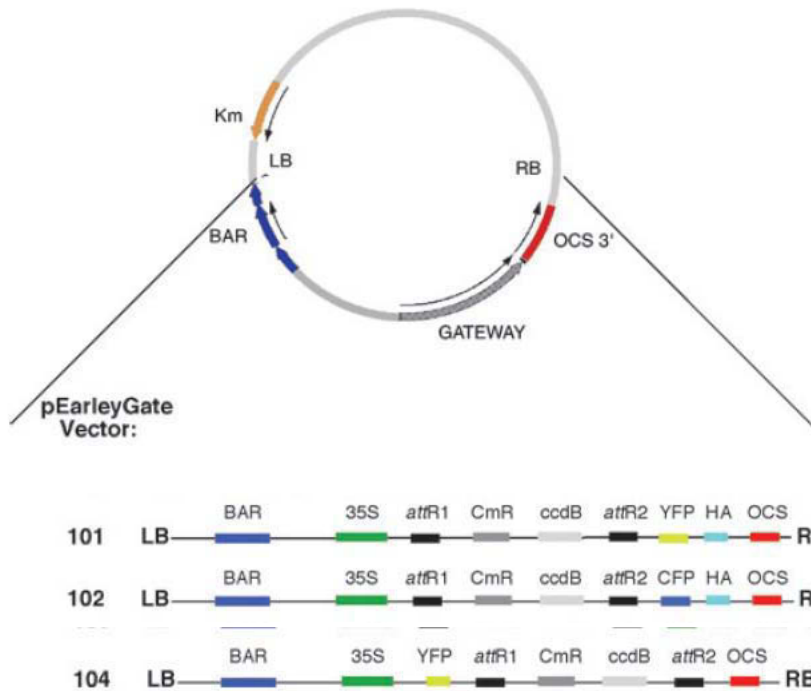
Comments for pDEST22
8930 nucleotides

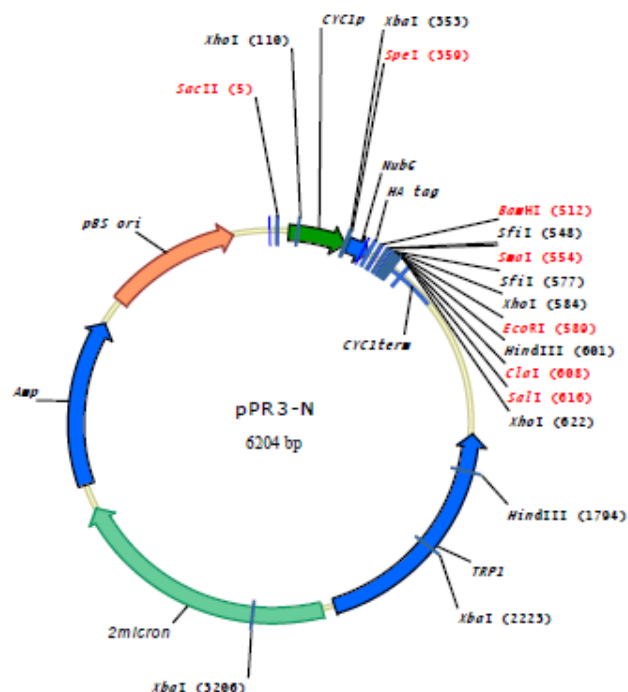
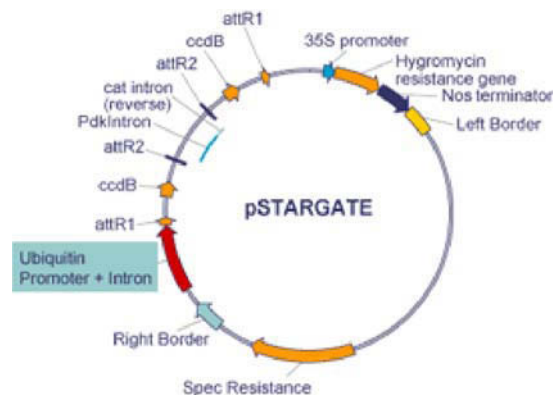
ADH1 promoter: bases 272-1726
Nuclear localization signal (NLS): bases 1734-1754
GAL4 DNA activation domain: bases 1761-2105
attR1 site: bases 2121-2245
Chloramphenicol resistance (Cm^R) gene: bases 2495-3154
ccdB gene: bases 3495-3800
attR2 site: bases 3841-3965
ADH1 transcription termination region: bases 4203-4360
f1 origin: bases 4687-5142
TRP1 gene: bases 5245-5919 (c)
ARS4/CEN6 origin: bases 6455-6972
Ampicillin (bla) resistance gene: bases 7104-7964
pUC origin: bases 8109-8782





Organisation des régions ADN-T des vecteurs pEarleyGate 101, 102 et 104 :





Sequence analysis of the NubG-HA reading frame:

Restriction sites: BamH I, Sfi I, Sma I, Sfi I.

Sequence: gct gga tcc aag cag tgg tat caa cgc aga gtg gcc att acg gcc cgg gaa aaa aca tgt cgg cgc cct cgg cct ctc gag

Protein sequence: A G S K Q W Y Q R R V A I T A R E K T C R P P R P L E

NubG-HA reading frame

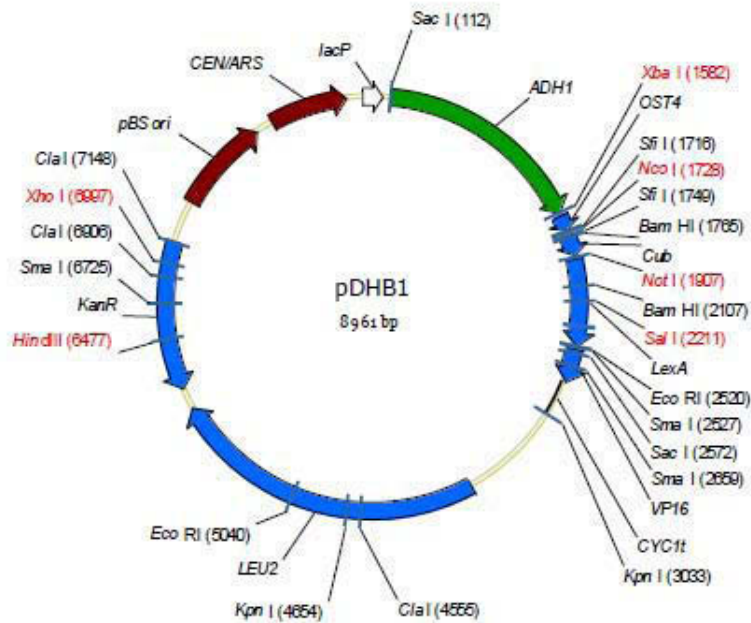
Restriction sites: EcoR I, Cla I, Sal I.

Sequence: aat tgg ata tca agc tta tgg ata cgg tgg acc

Protein sequence: N S I S S L S I P S T

Vector features

Position	Feature	Position	Feature
Start: 62 End: 352	CYC1 promoter	Start: 1508 End: 2791	TRP1 auxotrophic marker
Start: 364 End: 480	NubG, ubiquitin amino acids 1-38	Start: 2855 End: 4202	2micron origin of replication
Start: 481 End: 510	HA epitope tag	Start: 4335 End: 5193	AmpR resistance gene
Start: 622 End: 883	CYC1 terminator	Start: 5328 End: 5995	pBluescript origin of replication



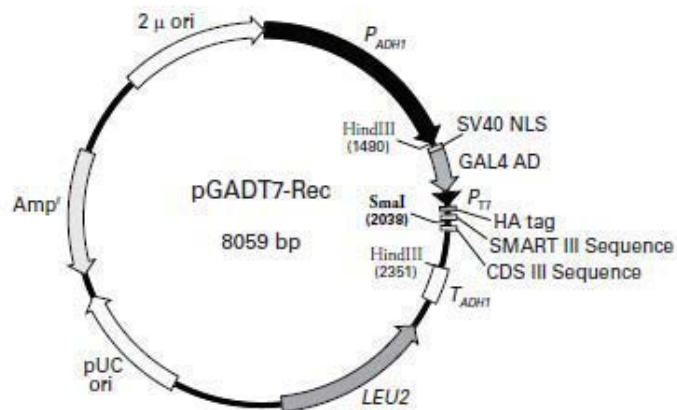
Sequence alignment showing restriction sites (Sfi I, Nco I, Sfi I) and open reading frames (ORFs) for Ost4 and Cub.

Ost4 ORF: cct aag aac gcg gcc att acg gcc agg cct cca tgg tat atc ttg gcc gcc tgc gcc atg tgc ggg
 P K N A A I T A R P P W Y I L A A S A M S G

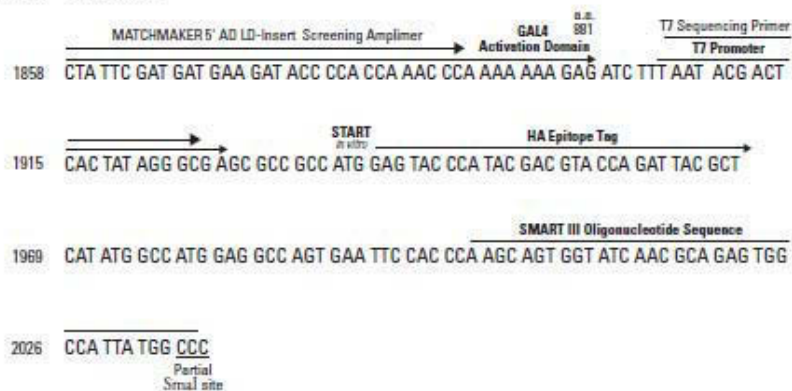
Cub ORF: cct aag aac gcg gcc att acg gcc agg cct cca tgg tat atc ttg gcc gcc tgc gcc atg tgc ggg
 P K N A A I T A R P P W Y I L A A S A M S G

Vector features

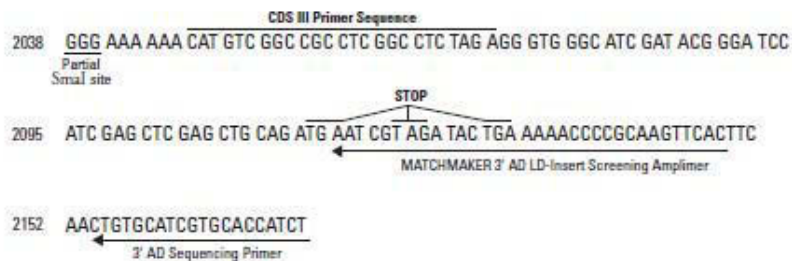
Position	Feature	Position	Feature
Start: 107 End: 1580	ADH1 promoter	Start: 3764 End: 5974	LEU2 auxotrophic marker
Start: 1598 End: 1705	Ost4	Start: 6123 End: 7156	KanR resistance gene
Start: 1754 End: 1888	Cub	Start: 7449 End: 8116	pBS replication origin (E. coli)
Start: 1913 End: 2518	LexA	Start: 8238 End: 8777	CEN/ARS replication origin (yeast)
Start: 2536 End: 2773	VP16	Start: 8890 End: 51	lacP
Start: 2775 End: 3031	CYC1 terminator		

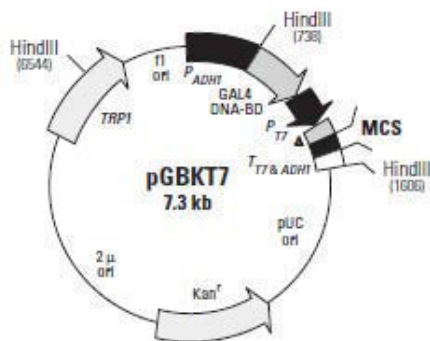


SMART™ III terminus

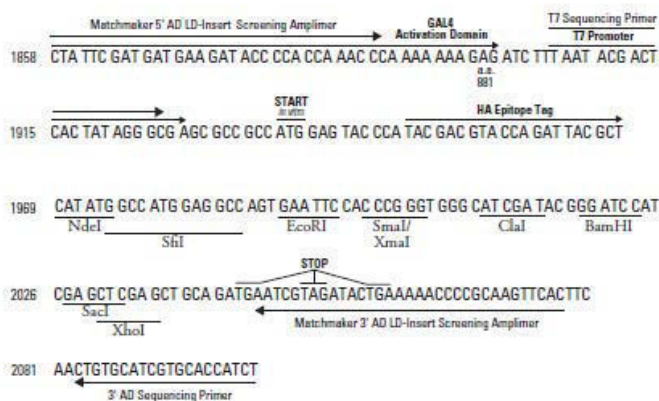
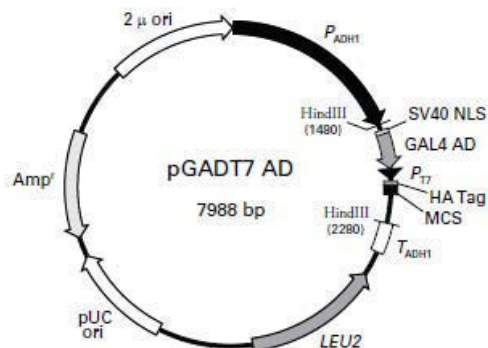
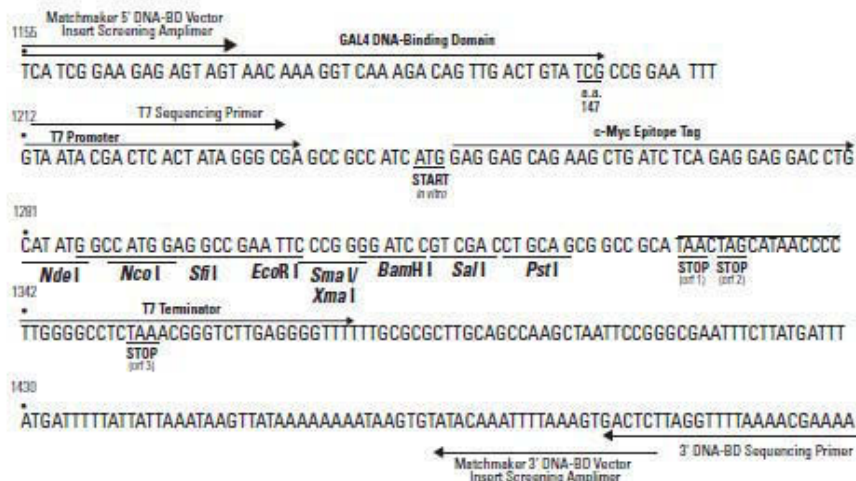


CDS III terminus





▲ c-Myc epitope tag



Annexe 4 : Liste des couples d'amorces

Nom du couple d'amorce	Utilisé pour	Sens/Antisens	Sequence	Température d'hybridation
TaGW2 FL CDS	TaGW2 CDS pleine-longueur	Sens	ATGGGGAACAGGATAGGGGGGA	61°C
		Antisens	CTACAACCATGCCAACCCTTG	
TaGW2-6A	Assignation chromosomique et recherche de marqueurs polymorphes	Sens	AGCGATTTGGCGTTCAATTC	Programme touchdown 64°C
		Antisens	GCCATCTGGTGAAC TTCCG	
TaGW2-6B		Sens	TGCCTTTTGAGCAACCAACGT	Programme touchdown 65°C
		Antisens	CAACAGGCAACGGTGGCG	
TaGW2-6B2		Sens	AGCGATTTGGCGTTCAATTC	Programme touchdown 65°C
		Antisens	GTCAGTAACAGGCAACGGTG	
TaGW2-6D		Sens	ATAGGAAATCCTGCTTGTGGGAGC	65°C
		Antisens	GTCAGCAAAAGGCAACGGTG	
AttbTaGW2-FL-CDS	Clonage Gateway du CDS pleine-longueur TaGW2	Sens	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGGGAAC AGGATAGGGGGGA	61°C
		Antisens	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTACAACCATG CCAACCCTTG	
PDI-intron	Absence de contamination ADN g	Sens	CTTATGCAATCCAATGCTGG	54°C
		Antisens	TCTATGGTTCTGAAGAGGACC	
A-SNPintron7 (T/C)	Génotypage du marqueur TaGW2_A_Y1061	Sens	(M13)-CAAATTATTGTCTCACCTTTG	61°C
		Antisens Allèle T	CTGCTAAGAGAGAGAATGTCAA	
		Antisens Allèle C	CTGCTAAGAGAGAGAATGTCAG	62°C
B-SNPintron7 (T/G)	Génotypage du marqueur TaGW2_B_Y447	Sens Allèle T	ATCAACTGCAAAATAAGGTT	58°C
		Sens Allèle G	ATCAACTGCAAAATAAGGTG	59°C
		Antisens	(M13)-TCAGGACAAGTAGATTGCG	
A-SNPR(-593) (A/G)	Génotypage du marqueur TaGW2_A_R(-593)	Sens Allèle A	GGTGAGACGAAAATAAATCGA	60°C
		Sens Allèle G	GGTGAGACGAAAATAAATCGG	62°C
		Antisens	(M13)-GTAACGCTTGATAAACATAGGTAAT	

TaGW2_318	Quantification de TaGW2 dans les tissus végétatifs	Sens	CCCCAACTATGCTGTGGAGT	58°C
		Antisens	GAATGGAAAGACCGCATGTT	
TaGW2-A	Quantification des transcrits homéologues par PCR quantitative	Sens	AGGAAATCCTTCTTGTGGGAGC	64°C
		Antisens	GCCATCTGGTGAACCTTCCG	
TaGW2-B		Sens	AATCCTGCTTGTGGGAGT	57°C
		Antisens	TGAACTCTCCACAGCACA	
TaGW2-D		Sens	ATAGGAAATCCTGCTTGTGGGAGC	66°C
		Antisens	GTCAGCAAAAGGCAACGGTG	
Ta2776	Gène de référence dans les analyses transcriptionnelles des homéologues	Sens	CGATTTCAGAGCAGCGTATTGTTG	61°C
		Antisens	AGTTGGTCGGGTCTCTTCTAAATG	
Ta54227	Gène de référence dans les lignées RNAi	Sens	CAAATACGCCATCAGGGAGAACATC	61°C
		Antisens	CGCTGCCGAAACCACGAGAC	
RNAiTaGW2 endo	TaGW2 endogène, toutes copies confondues dans les lignées RNAi	Sens	AACAGGATAGGGGGGAGGAGGAA	61°C
		Antisens	ATGGGGCACTCCTCCAGGTCA	
RNAi TaGW2-Aendo	TaGW2-A endogène dans les lignées RNAi	Sens	AACAGGATAGGGGGGAGGAGGAA	63°C
		Antisens	GCCATCTGGTGAACCTTCCG	
RNAi TaGW2-Bendo	TaGW2-B endogène dans les lignées RNAi	Sens	AACAGGATAGGGGGGAGGAGGAA	63°C
		Antisens	TGAACTCTCCACAGCACA	
RNAi TaGW2-Dendo	TaGW2-D endogène dans les lignées RNAi	Sens	AACAGGATAGGGGGGAGGAGGAA	65,5°C
		Antisens	GTCAGCAAAAGGCAACGGTG	
PDK-intron	Présence de la cassette RNAi	Sens	GTTAGAAATTCCAATCTGCT	54°C
		Antisens	ATGATGATGATGATAACTGC	
(1-39)TaGW2	Mutants de délétion TaGW2 pour les expériences de localisation subcellulaire	Sens	ATGGGGAACAGGATAGGGGGGA	64°C
		Antisens	GAGGATCAACTTGCGTAGCTTCTTC	
(19-39)TaGW2		Sens	ATGAGGCCGCAGGGGCTGTAC	64°C
		Antisens	GAGGATCAACTTGCGTAGCTTCTTC	

pGAD-insert screening	Séquençage de l'insert dans le vecteur pGADT7	Sens	CTATTCGATGATGAAGATACCCACCAAACCC	(Programme PCR en 2 temps : 95°C 30s ; 68°C, 3 min)
		Antisens	GTGAACCTGCGGGGTTTTTCAGTATCTACGATT	
pDHB1	Séquençage de l'insert dans le vecteur pDHB1	Sens	TTTCTGCACAATATTTCAAGC	52°C
		Antisens	GTAAGGTGGACTCCTTCT	
TaGW2 (1-320)	Mutant de délétion TaGW2 pour criblage de la banque ADNc	Sens	ATGGGGAACAGGATAGGGGGGA	63°C
		Antisens	CATGCTTCCACCAGCAATGTTG	
	Clonage de TaGW2(1- 320) dans le vecteur pGBKT7	Sens	CATGGAGGCCGAATTCGGGAACAGGATAGGGGGGAGGAGG	63°C
		Antisens	GCAGGTCGACGGATCCCATGCTTCCACCAGCAATGTTGCA	
SfiI-TaGW2	Clonage de <i>TaGW2</i> CDS dans le vecteur pDHB1	Sens	CGCAGAGTGGCCATTACGGCCATGGGGAACAGGATAGGGG GGAG	63°C
		Antisens	CGCAGAGTGGCCGAGGCGGCCAACCATGCCAACCCTTGCG AGTG	
pPR3-N	Séquençage de l'insert dans le vecteur pPR3-N	Sens	GTCGAAAATTCAAGACAAGG	56°C
		Antisens	AAGCGTGACATAACTAATTAC	

Annexe 5 : Protocole de précipitation d'ADN à l'éthanol

Transférer la solution d'ADN à précipiter dans un microtube. Le volume du microtube doit être au minimum 4 fois supérieur au volume de la solution d'ADN à précipiter. Ajouter l'équivalent d' $1/10^{\text{ème}}$ du volume de la solution d'ADN d'acétate de sodium 3M, pH 5,2. Ajouter ensuite 2 volumes d'éthanol absolu froid, inverser le tube, et placer à -20°C pendant au minimum 1h. Centrifuger l'échantillon pendant 15 min, à 14000 rpm, à 4°C. Retirer le surnageant, centrifuger de nouveau le tube et retirer le restant de surnageant. Ajouter 200 µl d'éthanol 70% (v/v) froid et centrifuger pendant 5 min, à 14000 rpm, à 4°C. Retirer le surnageant et resuspendre le culot dans un volume de tampon TE compris entre 10 et 40 µl.

Annexe 6 : Boîtes putatives de fixation de facteurs de transcription dans le promoteur de *TaGW2-D*. Un fragment de la séquence promotrice de *TaGW2-D* (2004 pb en amont du site d'initiation de la transcription) a été analysé à l'aide de l'outil en ligne PlantPan (<http://plantpan.mbc.nctu.edu.tw>). Les boîtes décrites dans la littérature dans la signalisation à l'ABA et au stress hydrique sont colorées en jaune.

Site	Motif	Espèce	Facteur de transcription reconnaissant le motif	Base de données source	Motif décrit entre autre dans :
612	CTGCTTCC	maïs	<u>ABI4</u>	JASPER	Signalisation à l'ABA
966	GTGCCCTT	maïs	<u>ABI4</u>	JASPER	
1793	GCGCCTTC	maïs	<u>ABI4</u>	JASPER	
1879	GCGCTCCC	maïs	<u>ABI4</u>	JASPER	
1203	ACGTG	Arabidopsis	<u>ABRELATERD1</u>	PLACE	Signalisation à l'ABA, stress hydrique
1672	TACGTGTC	Arabidopsis	<u>ABRE-like</u>	AGRIS	Signalisation à l'ABA, stress hydrique et froid
1673	ACGTGTC	riz/Arabidopsis	<u>M</u>	PLACE	Signalisation à l'ABA
634	ACACCTG	Arabidopsis	<u>DPBFCOREDCDC3</u>	PLACE	Signalisation à l'ABA par le FT ABI5; grain
658	ACACGAG	Arabidopsis	<u>DPBFCOREDCDC3</u>	PLACE	
1693	ACACAGG	Arabidopsis	<u>DPBFCOREDCDC3</u>	PLACE	
28	CATTTG	Arabidopsis	<u>MYCCONSUSAT</u>	PLACE	Signalisation à l'ABA par les FT de type MYC, stress au froid (fixation du FT ICE1)
110	CATATG	Arabidopsis	<u>MYCCONSUSAT</u>	PLACE	
230	CAATTG	Arabidopsis	<u>MYCCONSUSAT</u>	PLACE	
256	CATTTG	Arabidopsis	<u>MYCCONSUSAT</u>	PLACE	
431	CAGCTG	Arabidopsis	<u>MYCCONSUSAT</u>	PLACE	
635	CACCTG	Arabidopsis	<u>MYCCONSUSAT</u>	PLACE	
1387	CAAGTG	Arabidopsis	<u>MYCCONSUSAT</u>	PLACE	
1481	CACTTG	Arabidopsis	<u>MYCCONSUSAT</u>	PLACE	
260	TGTAAAG	maïs/blé/orge/tabac	<u>-300CORE</u>	PLACE	grain, albumen
100	TGCAAAGT	blé	<u>-300ELEMENT</u>	PLACE	grain, albumen
100	TGCAAAG	riz	<u>PROLAMINBOXOSGLUB1</u>	PLACE	grain, albumen
1282	AACAAAC	riz	<u>AACACOREOSGLUB1</u>	PLACE	grain, albumen

Site	Motif	Espèce	Facteur de transcription reconnaissant le motif	Base de données source	Motif décrit entre autre dans :
979	ACGT	Arabidopsis	<u>ACGTATERD1</u>	PLACE	Stress hybrique
1474	ACGT	Arabidopsis	<u>ACGTATERD1</u>	PLACE	
48	AAACCA	Arabidopsis	<u>MYB1AT</u>	PLACE	MYB, stress hydrique
154	CTGTTA	Arabidopsis/petunia	<u>MYBCORE</u>	PLACE	Stress hydrique, Fixation de FT de type MYB
469	CAGTTA	Arabidopsis/petunia	<u>MYBCORE</u>	PLACE	
878	ctagCCAAAatagatcaa	Arabidopsis	<u>AG</u>	TRANSFAC	(Dans promoteur de gènes homéotiques, développement floral)
1456	CCTTTAAAGT	Arabidopsis	<u>Agamous</u>	JASPER	
499	ttagggcATGATtcgatcc	Arabidopsis	<u>ATHB-9</u>	TRANSFAC	(Dans promoteur de gènes homéotiques)
1616	AAATTAAT	Arabidopsis	<u>Bellringer</u>	AGRIS	(Dans promoteur de gènes homéotiques, développement floral)
13	ATTA	Arabidopsis	<u>Core</u>	AGRIS	(Dans promoteur de gènes homéotiques)
298	ATTA	Arabidopsis	<u>Core</u>	AGRIS	
498	ATTA	Arabidopsis	<u>Core</u>	AGRIS	
761	ATTA	Arabidopsis	<u>Core</u>	AGRIS	
851	ATTA	Arabidopsis	<u>Core</u>	AGRIS	
1559	ATTA	Arabidopsis	<u>Core</u>	AGRIS	
1618	ATTA	Arabidopsis	<u>Core</u>	AGRIS	
1622	ATTA	Arabidopsis	<u>Core</u>	AGRIS	
1278	tcgCAACAaacc	Arabidopsis	<u>RAV1</u>	TRANSFAC	(Dans promoteur de gènes homéotiques, développement floral)
1281	CAACA	Arabidopsis	<u>RAV1-A</u>	AGRIS	
635	CACCTG	Arabidopsis	<u>RAV1-B</u>	AGRIS	
635	CACCTG	Arabidopsis	<u>RAV1BAT</u>	PLACE	
575	cgatgtcCCCGAgc	Arabidopsis	<u>ANT</u>	TRANSFAC	Division et élongation cellulaire (fixation du FT AINTEGUMENTA)
1823	ctccacCCCGAgg	Arabidopsis	<u>ANT</u>	TRANSFAC	
446	gttTCAGCaac	Arabidopsis	<u>CDC5</u>	TRANSFAC	Contrôle du cycle cellulaire
894	aaaTCAGCaag	Arabidopsis	<u>CDC5</u>	TRANSFAC	
1004	ctaTCAGCgag	Arabidopsis	<u>CDC5</u>	TRANSFAC	
806	AGATAATA	Arabidopsis	<u>ARR10</u>	JASPER	(fixation de FT de type MYB)
1045	AGATCATG	Arabidopsis	<u>ARR10</u>	JASPER	
1702	AGATTTTC	Arabidopsis	<u>ARR10</u>	JASPER	
1992	AGATCTGG	Arabidopsis	<u>ARR10</u>	JASPER	

Site	Motif	Espèce	Facteur de transcription reconnaissant le motif	Base de données source	Motif décrit entre autre dans :
407	TGATT	Arabidopsis	<u>ARR1AT</u>	PLACE	(fixation des FT de type MYB ARR1 et ARR2)
414	CGATT	Arabidopsis	<u>ARR1AT</u>	PLACE	
507	TGATT	Arabidopsis	<u>ARR1AT</u>	PLACE	
524	TGATT	Arabidopsis	<u>ARR1AT</u>	PLACE	
693	CGATT	Arabidopsis	<u>ARR1AT</u>	PLACE	
849	TGATT	Arabidopsis	<u>ARR1AT</u>	PLACE	
864	TGATT	Arabidopsis	<u>ARR1AT</u>	PLACE	
1105	GGATT	Arabidopsis	<u>ARR1AT</u>	PLACE	
453	CAACCaag	orge	<u>GAmYb</u>	TRANSFAC	Signalisation à l'acide gibbérellique par les FT MYB
1053	AGCAACCGTC	orge	<u>GAMYB</u>	JASPER	
1055	CAACCgtc	orge	<u>GAmYb</u>	TRANSFAC	
1425	AATAACCGTA	orge	<u>GAMYB</u>	JASPER	
1282	AACAAAC	Arabidopsis	<u>MYB4</u>	AGRIS	(Fixation de FT de type MYB)
604	CCTACC	maïs	<u>MYBPZM</u>	PLACE	(Fixation de FT de type MYB, grain)
255	TCATTgtaaagcccc	maïs	<u>O2</u>	TRANSFAC	Fixation du FT opaque2, grain
296	TCATTaggaaattgct	maïs	<u>O2</u>	TRANSFAC	
759	TCATTaaaaatgttac	maïs	<u>O2</u>	TRANSFAC	
1594	GGATATGC	Arabidopsis/Medicago/orge	<u>P1BS</u>	PLACE	(Fixation de FT de type MYB)
286	TGAC	riz	<u>WRKY71OS</u>	PLACE	Signalisation à l'acide gibbérellique par FT de type MYB
1212	TGAC	riz	<u>WRKY71OS</u>	PLACE	
1647	TGAC	riz	<u>WRKY71OS</u>	PLACE	
663	agtAAAAGaac	maïs	<u>PBF</u>	TRANSFAC	Fixation du FT PBF-1 (type DOF)
751	tttAAAAGtca	maïs	<u>PBF</u>	TRANSFAC	
1458	tttAAAAGtag	maïs	<u>PBF</u>	TRANSFAC	
480	AAAGA	maïs	<u>MNB1A</u>	JASPER	(Fixation de FT de type DOF)
667	AAAGA	maïs	<u>MNB1A</u>	JASPER	
1491	AAAGA	maïs	<u>MNB1A</u>	JASPER	

Site	Motif	Espèce	Facteur de transcription reconnaissant le motif	Base de données source	Motif décrit entre autre dans :
259	ttgTAAAGccc	maïs	<u>Dof1</u>	TRANSFAC	(Fixation de FT de type DOF)
1430	ccgTAAAGtga	maïs	<u>Dof1</u>	TRANSFAC	
263	AAAGCC	maïs	<u>Dof2</u>	JASPER	
1436	agtgAAAGCta	maïs	<u>Dof2</u>	TRANSFAC	
1440	AAAGCT	maïs	<u>Dof2</u>	JASPER	
103	AAAGTA	maïs	<u>Dof3</u>	JASPER	
755	AAAGTC	maïs	<u>Dof3</u>	JASPER	
927	AAAGTG	maïs	<u>Dof3</u>	JASPER	
1462	AAAGTA	maïs	<u>Dof3</u>	JASPER	
334	AAAG	maïs	<u>DOFCOREZM</u>	PLACE	
1169	AAAG	maïs	<u>DOFCOREZM</u>	PLACE	
1327	AAAG	maïs	<u>DOFCOREZM</u>	PLACE	
1374	AAAG	maïs	<u>DOFCOREZM</u>	PLACE	
253	TGTCA	riz	<u>BIHD1OS</u>	PLACE	Défense contre des pathogènes
1676	TGTCA	riz	<u>BIHD1OS</u>	PLACE	
118	atCTACta	maïs	<u>P</u>	TRANSFAC	(Dans promoteur de gènes de synthèse des flavonoïdes)
603	tcCTACtt	maïs	<u>P</u>	TRANSFAC	
1442	agCTACCaa	maïs	<u>P</u>	TRANSFAC	